Using Semantic Technologies to Resolve Heterogeneity Issues in Sustainability and Disaster Management Knowledge Bases

A thesis submitted for the degree of
Doctor of Philosophy

Lida Ghahremanlou
B.CS, M.CS

School of Computer Science and Information Technology,
College of Science, Engineering, and Health,
RMIT University,
Melbourne, Victoria, Australia.

September, 2014
Declaration

I certify that except where due acknowledgement has been made, the work is that of the author alone; the work has not been submitted previously, in whole or in part, to qualify for any other academic award; the content of the thesis is the result of work which has been carried out since the official commencement date of the approved research program; any editorial work, paid or unpaid, carried out by a third party is acknowledged; and, ethics procedures and guidelines have been followed.

Lida Ghahremanlou
School of Computer Science and Information Technology
RMIT University
June 28, 2015
Acknowledgments

First and foremost, I would like to express my heartfelt gratitude to my supervisors, Associate Professor James Thom and Dr. Liam Magee, for their continuous guidance, encouragement and patience throughout this research project. I could not be prouder of my academic roots and research values that they have given to me.

Special thanks goes to, Dr. Liam Magee, for his support and guidance in developing the Circles of Sustainability software user interface. I would not have be able to complete my experiments without his timely assistance in implementing the technical components as a part of the project development.

This research is co-funded by the Australian Research Council (ARC), a linkage project LP0990509 on Accounting for Sustainability: Developing an Integrated Approach for Sustainability Assessments, and I would like to thank Professor Paul James for providing me the opportunity to work with him and his wonderful team. To my colleagues at the Global Compact Cities Programme, Elizabeth Ryan, Sandra Moye, Julia Laidlaw and Cynthia Lam, thank you for your great support during the user-study experiment.

My other affiliation is with IBM Research – Australia through a two-year internship program since January 2012, supported by the Australian Computer Society (ACS). I would like to express my gratitude to Dr. Jürg von Känel, the leading director of the IBM Research – Australia, who is my mentor at the start of my career. I am thankful for the opportunity to learn and enrich my research skills at IBM.

To the staff and students at RMIT University, the School of Computer Science and Information Technology, I am grateful for the chance to be a part of this research group.

None of this would have been possible without support from my parents, Maryam Gharaei and Parviz Ghahremanlou, who instil within me a love of creative pursuits, science and language, all of which finds a place in this thesis. Thank you for giving me the courage to widen my horizons. My siblings, Lina and Hamed, you have also been the best of friends along this journey and I could not have been where I am without your unconditional love.
Credits

Portions of the material in this thesis have previously appeared in the following publications:


Portions of the material in this thesis have been submitted for publications:


- Butler, C. J., Sherchan, W., Ghahremanlou L., Han, B., Pervin, S. and Lai, J. Harnessing Twitter and Instagram for Disaster Management: Project ACT.

The thesis was written in the TeXstudio editor on Ubuntu 12.4 and Mac OS X 10.7.2, and typeset using the LATEX document preparation system.

All trademarks are the property of their respective owners.
Contents

Abstract 1

1 Introduction 3
  1.1 Motivation and Solutions ........................................... 6
  1.2 Research Questions .................................................. 8
  1.3 Research Contributions ............................................ 10
  1.4 Thesis Overview ..................................................... 10

2 Knowledge Bases and Ontologies 12
  2.1 Semantic Heterogeneity .............................................. 12
  2.2 A History of Development of Knowledge Base Systems .......... 15
    2.2.1 Transferring Process .......................................... 15
    2.2.2 Modelling Process ........................................... 15
  2.3 Ontologies ........................................................... 17
    2.3.1 Ontology Components .......................................... 18
    2.3.2 Taxonomies ..................................................... 19
    2.3.3 Ontology Specification Languages ......................... 20
      Web Ontology Language ........................................... 20
      Resource Description Framework ................................ 21
  2.4 Ontology Granularity .............................................. 22
    2.4.1 Upper Ontologies ............................................. 23
    2.4.2 Domain Ontologies ............................................ 23
2.4.3 Application Ontologies ........................................ 24
2.5 Ontology Learning Approaches .................................. 24
2.6 Ontology Engineering Methodologies ......................... 25
  2.6.1 Skeletal Methodology ......................................... 25
  2.6.2 Formal Method for Ontology Engineering ................. 25
  2.6.3 On-To-Knowledge Methodology ............................... 26
  2.6.4 SENSUS-based Ontology Methodology ................. 27
  2.6.5 METHONTOLOGY ............................................ 27
2.7 Ontology Design Patterns ......................................... 29
2.8 Ontology Evaluation Methodologies ............................ 32
  2.8.1 OntoClean .................................................. 32
  2.8.2 OntoMetric ................................................ 33
  2.8.3 ROMEO ................................................... 34
2.9 Ontologies used in Sustainability and Sustainability Reporting Indicator Sets 35
  2.9.1 Taxonomies of Sustainability Reporting Indicator Sets .... 37
2.10 Knowledge bases and Ontologies used in Social Media Content Analysis .... 39
  2.10.1 Using Probabilistic Language Models ...................... 41
  2.10.2 Using Ontologies ......................................... 42
2.11 Discussion ...................................................... 44
2.12 Summary ........................................................ 47

3 Research Design .................................................... 48
  3.1 A Knowledge Base for Representing Heterogeneity of Sustainability Indicator Sets ................................................................. 49
    3.1.1 Adapting an Ontology Engineering Approach ............ 49
    3.1.2 Using Sustainability Indicator Sets ................... 50
    3.1.3 Using an Ontology Design Pattern ..................... 50
    3.1.4 Developing Two Ontology Design Candidates .. .... 51
    3.1.5 Applying an Ontology Evaluation Method ............ 52
  3.2 Mechanisms for Browsing Heterogeneous Sustainability Indicators .... 53
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2.1</td>
<td>Structure of Circles of Sustainability Framework</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>Circles of Sustainability Model</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>Circles of Sustainability Methodology</td>
<td>56</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Developing Circles of Sustainability Software</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>Database</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>SPARQL-end</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>Back-end</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>Play Framework</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>Ontology Model</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>Front-end</td>
<td>60</td>
</tr>
<tr>
<td>3.2.3</td>
<td>Browsing Heterogeneous Sustainability Indicator Sets</td>
<td>60</td>
</tr>
<tr>
<td>3.2.4</td>
<td>Evaluating User Interfaces</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td>Usability Measures</td>
<td>62</td>
</tr>
<tr>
<td>3.2.5</td>
<td>Conducting a User Study</td>
<td>63</td>
</tr>
<tr>
<td>3.3</td>
<td>Resolving Heterogeneity of Georeferencing Emergency Tweets</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>Developing OzCrisisTracker Geotagger Module</td>
<td>64</td>
</tr>
<tr>
<td></td>
<td>Processing Geographic References</td>
<td>64</td>
</tr>
<tr>
<td></td>
<td>Toponym Recognition</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>Applying Tweet Toponym Recognition</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td>Toponym Resolution</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td>Applying an Ontological Model</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td>Using Geographical Knowledge Bases</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>Evaluating the OzCT Geotagger</td>
<td>67</td>
</tr>
<tr>
<td>3.4</td>
<td>Summary</td>
<td>68</td>
</tr>
<tr>
<td>4</td>
<td>An Ontology for Sustainability Indicator Sets</td>
<td>69</td>
</tr>
<tr>
<td>4.1</td>
<td>OSIS Pre-Design Process</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>Specification</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>Knowledge Acquisition</td>
<td>73</td>
</tr>
<tr>
<td>4.2</td>
<td>OSIS Design Process</td>
<td>75</td>
</tr>
</tbody>
</table>
## CONTENTS

5.2.3 Hypotheses ................................. 106  
5.2.4 Questionnaire ............................. 107  
5.2.5 Characteristics of Target Audiences .......... 108  
  Sample Size of Participants ..................... 109  
5.2.6 Statistical Tests ............................ 109  
5.2.7 Tasks .................................... 110  
  Scripts ..................................... 110  
5.3 Results and Discussion ........................ 111  
  5.3.1 Quantitative Results ...................... 111  
    Background Knowledge of the Participants .... 112  
    Overall Usefulness ......................... 113  
    Ease of Use of Browsing Indicator Sets ........ 115  
    Which View is Preferable? .................... 119  
  5.3.2 Qualitative Results ...................... 119  
    Software Strengths ........................ 120  
    Software Limitations ...................... 120  
    Suggestions for Improvements ............... 122  
5.4 Summary .................................... 123  

6 Geotagging Tweets in the Context of Disaster Management 125  
  6.1 Data Collection .............................. 126  
  6.2 Tweet Toponym Recognition .................. 129  
  6.3 Tweet Toponym Resolution .................... 130  
  6.4 Evaluation .................................. 134  
    6.4.1 Evaluation against Human Judgement (Ground Truth) .... 134  
      Evaluation Measures ....................... 135  
    6.4.2 Comparison with Existing GeoCoding Platforms ........ 135  
      Evaluation Measures ....................... 136  
  6.5 Results and Discussion ..................... 137  
    6.5.1 Evaluation Results ...................... 137
## Contents

Time Analysis .................................................. 140

6.5.2 Ontological Toponym Resolution .......................... 142

Complexity of SPARQL Queries ................................. 144

6.6 Summary ................................................... 146

7 Conclusions and Future Work ................................. 147

7.1 Solutions .................................................. 147

7.2 Limitations ............................................... 149

7.3 Future Work ............................................... 150

7.4 Final Remarks ............................................ 152

Appendix A OWL Code of OSIS Designs ....................... 153

Appendix B CoS User Study - Ethics .......................... 160

Appendix C CoS User Study - Questionnaire ................. 165

Appendix D CoS User Study - Scripts ......................... 171

Appendix E CoS User Study - Qualitative Responses ....... 180

Appendix F Glossary ........................................... 185

Bibliography .................................................... 188
List of Figures

1.1 Our Model to Resolve Semantic Heterogeneity Issues of Two Domains .......................... 5
1.2 Overview of Three Research Questions and Users Benefits of the Thesis Outcomes ........... 9

2.1 Levels of Granularity of Ontologies Used in this Thesis ................................. 22
2.2 METHONTOLOGY [Gómez-Pérez et al., 2004] Overview Adapted from [Corcho et al., 2005] .................................................. 28
2.3 ROMEO Methodology Adapted from [Yu et al., 2009] ........................................ 34
2.4 A Snapshot of Sustainability GRI Indicator Set - Not all categories are shown .......... 37
2.5 A Snapshot of OECD Sustainability Indicator Set - Not all categories are shown .......... 38

3.1 Circles of Sustainability Framework (domains and subdomains) [Magee et al., 2013] ........ 55
3.2 Model-View-Controller Pattern ................................................................................. 58

4.1 OSIS Ontology Engineering Process ........................................................................ 71
4.2 A Sample GRI Indicator ......................................................................................... 73
4.3 A Sample OECD Indicator ..................................................................................... 74
4.4 Value Partition - Pattern 1 ..................................................................................... 78
4.5 UML Diagram of SOSIS Design Using Value Partition Pattern 1 ......................... 79
4.6 Value Partition - Pattern 2 ..................................................................................... 80
4.7 UML Diagram of GOSIS Design Using Value Partition Pattern 2 ......................... 81
4.8 Exampe of RDF Triples ......................................................................................... 83
4.9 F-measure of Two Ontology Design Candidates, GRI and OECD are used in constructing the ontologies and UN is not used ............................................ 95
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Login Page of CoS Software</td>
<td>98</td>
</tr>
<tr>
<td>5.2</td>
<td>Screenshot of Tabular View</td>
<td>102</td>
</tr>
<tr>
<td>5.3</td>
<td>Screenshot of Circular View</td>
<td>103</td>
</tr>
<tr>
<td>5.4</td>
<td>Summary of the Background Knowledge of the Participants</td>
<td>111</td>
</tr>
<tr>
<td>6.1</td>
<td>OzCT Geotagger Structure</td>
<td>128</td>
</tr>
<tr>
<td>6.2</td>
<td>Geographical Focus of Three Geotaggers</td>
<td>139</td>
</tr>
<tr>
<td>6.3</td>
<td>Precision and Accuracy of Geographical Focus of OzCT Geotagger</td>
<td>140</td>
</tr>
<tr>
<td>6.4</td>
<td>Google Map Representation of the Distance between Two Streets as an Example of Ontological Toponym Resolution</td>
<td>145</td>
</tr>
<tr>
<td>B.1</td>
<td>Ethics Approval for CoS Research</td>
<td>161</td>
</tr>
<tr>
<td>B.2</td>
<td>Plain Language Statement for the CoS User Study Evaluation - Page 1</td>
<td>162</td>
</tr>
<tr>
<td>B.3</td>
<td>Plain Language Statement for the CoS User Study Evaluation - Page 2</td>
<td>163</td>
</tr>
<tr>
<td>B.4</td>
<td>Plain Language Statement for the CoS User Study Evaluation - Page 3</td>
<td>164</td>
</tr>
<tr>
<td>C.1</td>
<td>Consent Form</td>
<td>166</td>
</tr>
<tr>
<td>C.2</td>
<td>Section 1 - Background Knowledge of the Participants - Page 1</td>
<td>167</td>
</tr>
<tr>
<td>C.3</td>
<td>Section 1 - Background Knowledge of the Participants - Page 2</td>
<td>168</td>
</tr>
<tr>
<td>C.4</td>
<td>Section 2 - Overall Usefulness of the CoS Software</td>
<td>169</td>
</tr>
<tr>
<td>C.5</td>
<td>Section 3 - Ease of Use of Browsing Mechanisms of Indicator Sets</td>
<td>170</td>
</tr>
<tr>
<td>D.1</td>
<td>CoS User Study - Script A - Page 1</td>
<td>172</td>
</tr>
<tr>
<td>D.2</td>
<td>CoS User Study - Script A - Page 2</td>
<td>173</td>
</tr>
<tr>
<td>D.3</td>
<td>CoS User Study - Script A - Page 3</td>
<td>174</td>
</tr>
<tr>
<td>D.4</td>
<td>CoS User Study - Script A - Page 4</td>
<td>175</td>
</tr>
<tr>
<td>D.5</td>
<td>CoS User Study - Script B - Page 1</td>
<td>176</td>
</tr>
<tr>
<td>D.6</td>
<td>CoS User Study - Script B - Page 2</td>
<td>177</td>
</tr>
<tr>
<td>D.7</td>
<td>CoS User Study - Script B - Page 3</td>
<td>178</td>
</tr>
<tr>
<td>D.8</td>
<td>CoS User Study - Script B - Page 4</td>
<td>179</td>
</tr>
<tr>
<td>E.1</td>
<td>Participants’ Responses to Q12</td>
<td>181</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

E.2 Participants’ Responses to Q25 - Part 1 ........................................ 182
E.3 Participants’ Responses to Q25 - Part 2 ........................................ 183
E.4 Participants’ Responses to Generat Statement ............................... 184
List of Tables

1.1 Shifting Emphasis on Disaster Management [Salter, 1997] ............................................. 4

3.1 Advantages versus Disadvantages of Value Partition Pattern 1 and Pattern 2 ............ 52

4.1 Ontology Requirements Specification Document for OSIS .................................................. 72

4.2 GOSIS design .......................................................................................................................... 85

4.3 Number of Sub-category and Indicators per Category for GRI and OECD Taxonomies 86

4.4 OSIS Role ............................................................................................................................... 87

4.5 OSIS Requirements 1 and 2 ...................................................................................................... 88

4.6 An Example of GRI Indicator (EN2) in XBRL format ......................................................... 90

4.7 An Example of GRI Indicator (EN2) in NTriples format ..................................................... 91

4.8 An Example of OECD Indicator in XML format ..................................................................... 93

4.9 Precision, Recall and F-measure of OSIS design Candidates against Seen and Unseen Frames of Reference ........................................................................................................... 94

5.1 Mapping between Variables and Questions ........................................................................... 106

5.2 Descriptive Statistics of Six Questions - Perceived Usefulness ........................................... 113

5.3 Spearman rho Correlations for Six Questions - Perceived Usefulness ................................. 114

5.4 Signed Wilcoxon Descriptive Statistics .................................................................................. 115

5.5 Signed Wilcoxon Test - Significant Difference and Z-value - Perceived Ease of Use ........ 116

5.6 Signed Wilcoxon Test - Perceived Ease of Use ..................................................................... 117

5.7 Users’ Insight of Working with Two Views ............................................................................. 118
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>Training (Jan-Feb) and Testing (Mar-Apr) Datasets Collected by the OzCrisisTracker (Retweets are Filtered)</td>
</tr>
<tr>
<td>6.2</td>
<td>A Set of Heuristics Developed from Training Dataset Used in OzCT Toponym Resolution</td>
</tr>
<tr>
<td>6.3</td>
<td>Ontology-Based Semantic Annotation Adapted [Bontcheva and Rout, 2012] and Compared with the OzCT Geotagger</td>
</tr>
<tr>
<td>6.4</td>
<td>Ground Truth versus OzCT Geotagger Outcome (500 Tweets)</td>
</tr>
<tr>
<td>6.5</td>
<td>Accuracy of the Geographical Focus of Definite Locations in Three Geotaggers (Confusion Matrix)</td>
</tr>
<tr>
<td>6.6</td>
<td>Geotagging Time of Three Geotagging Systems</td>
</tr>
<tr>
<td>6.7</td>
<td>Recall, Precision and F1-score for OzCT Geotagger</td>
</tr>
<tr>
<td>6.8</td>
<td>A Snapshot of OWL representation of a Tweet using Ontology Model</td>
</tr>
<tr>
<td>A.1</td>
<td>GOSIS Design Ontology OWL Code - Part 1</td>
</tr>
<tr>
<td>A.2</td>
<td>GOSIS Design Ontology OWL Code - Part 2</td>
</tr>
<tr>
<td>A.3</td>
<td>GOSIS Design Ontology OWL Code - Part 3</td>
</tr>
<tr>
<td>A.4</td>
<td>SOSIS Design Ontology OWL Code - Part 1</td>
</tr>
<tr>
<td>A.5</td>
<td>SOSIS Design Ontology OWL Code - Part 2</td>
</tr>
<tr>
<td>A.6</td>
<td>SOSIS Design Ontology OWL Code - Part 3</td>
</tr>
</tbody>
</table>
Abstract

This thesis examines issues of semantic heterogeneity in the domains of sustainability indicators and disaster management. We propose a model that links two domains with the following logic. While disaster management implies a proper and efficient response to a risk that has materialised as a disaster, sustainability can be defined as the preparedness to unexpected situations by applying measurements such as sustainability indicators. As a step to this direction, we investigate how semantic technologies can tackle the issues of heterogeneity in the aforementioned domains. In particular, we develop ontologies and browsing mechanisms for representing heterogeneous sustainability indicator sets. In addition, with the use of new and existing knowledge bases, we resolve the heterogeneity of georeferences of emergency tweets.

First, we consider approaches to resolve the heterogeneity issues of representing the key concepts of sustainability indicator sets. To develop a knowledge base, we apply the METHONTOLGY approach to guide the construction of two ontology design candidates: generic and specific. Of the two, the generic design is more abstract, with fewer classes and properties. Documents describing two indicator systems – the Global Reporting Initiative and the Organisation for Economic Co-operation and Development – are used in the design of both candidate ontologies. We then evaluate both ontology designs using the ROMEO approach, to calculate their level of coverage against the seen indicators, as well as against an unseen third indicator set (the United Nations Statistics Division). We also show that use of existing structured approaches like METHONTOLGY and ROMEO can reduce ambiguity in ontology design and evaluation for domain-level ontologies. It is concluded that where an ontology needs to be designed for both seen and unseen indicator systems, a generic and
reusable design is preferable.

Second, having addressed the heterogeneity issues at the data level of sustainability indicators in the first phase of the research, we then develop a software for a sustainability reporting framework – **Circles of Sustainability** – which provides two mechanisms for browsing heterogeneous sustainability indicator sets: a **Tabular** view and a **Circular** view. In particular, the generic design of ontology developed during the first phase of the research is applied to this software. Next, we evaluate the overall usefulness and ease of use for the presented software and the associated user interfaces by conducting a user study. The analysis of quantitative and qualitative results of the user study concludes that the **Circular** view is the preferred interface by most participants for browsing semantic heterogeneous indicators. Its satisfactory aspects include comparative visual experience of assessment process, better graphical display, ease of navigation and visual communication. The **Tabular** view, however, has advantages of simplicity and easy search.

Third, in the context of disaster management, we present a geotagger method for the **OzCrisisTracker** application that automatically detects and disambiguates the heterogeneity of georeferences mentioned in the tweets’ content with three possibilities: **definite**, **ambiguous** and **no-location**. Our method also semantically annotates the tweet components utilising existing and new ontologies. Experiments demonstrate that the precision and recall for detection of the definite locations are about 80%. We also concluded that the accuracy of geographic focus of our geotagger is considerably higher than other systems.

From a more general perspective the research contributions can be articulated as follows. The overall objective of this research is exploring approaches to tackle various problems of semantic heterogeneity for two related areas of sustainability indicator sets and disaster management at the data and user interface levels of sustainability indicators and georeferences of tweets. The knowledge bases developed in this research have been applied to the two domain applications. The thesis therefore demonstrates how semantic technologies, such as ontology design patterns, browsing tools and geocoding, can untangle data representation and navigation issues of semantic heterogeneity in sustainability and disaster management domains.
Chapter 1

Introduction

We know very little, and yet it is astonishing that we know so much, and still more astonishing that so little knowledge can give us so much power.

–Bertrand Russell

This research proposes the use of semantic technologies to resolve the heterogeneity issues of data representation for two related domains of sustainability and disaster management.

A common definition of the term sustainability is to ensure our use of resources is sufficient to meet the needs of the present without compromising the ability of future generations to meet their own needs [Burton, 1987]. Over the past 30 years, sustainability and sustainable development – not only in environmental, but also in economic and social areas – has become a key theme in both academic and popular literature. Since the publication of the Club of Rome’s The Limits to Growth in the early 1970s [Meadows et al., 2004], the sustainability of natural and social systems has become a pressing concern. This problem has become increasingly urgent, given the depletion of our natural stocks due to past and present economic development. To respond to such challenges, it is vitally important to measure the status of the complex economic, environmental and social systems. These provide the tools to manage current and future development responsibly.

In response to measuring and maintaining the sustainability status of a generic system, several indicator systems have also been developed and are in use today. Sustainability
indicators are typically numerical measures that provide key elements about physical, social or economic systems. In particular, they are used to show complex trends and cause-and-effect relationships rather than presenting simple data [Kuik and Gilbert, 2002]. There have been few efforts to represent multiple indicators formally, in spite of the fact that comparison of indicators and measurements across reporting contexts is a critical task. For example, GRI (See Figure 2.4) presents two sustainability indicators for the category of “Environmental Performance” and sub-category of “Material” including: EN1: Materials used by weight or volume, and EN2: Percentage of materials used that are recycle input materials, which are defined to measure materials used for different purposes.

Disaster management is also expected to make policies and activities to address the concerns such as biodiversity protection, sustainable use of land and water resources, greenhouse prevention gas emissions and pollution. According to Salter [1997], a shift of concerns has occurred in disaster management topics shown in Table 1.1. The example of uncertainties in this context are natural hazards and environmental risks, which are constant threats to human societies and their resilience.

During times of crisis, microblogging platforms such as Twitter have played an important role as a communication channel to distribute information. In particular, emergency tweets are valuable resources when tagged with their location information which could result in detecting unexpected events. However, one of the main challenges is resolving heterogeneous ambiguity in georeferencing tweet content, which has not been addressed sufficiently in the literature.

Related to both sustainability and natural disaster is the concept of Resilience, that
was first defined as the ability to maintain a steady state for an ecological system [Holling and Sanderson, 1996]. This definition has shifted over time. Adger et al. [2005] explain this shift in which the traditional perspective in the concept of resilience is an attempt to control changes in systems that are assumed to be stable, whereas, the modern view aims at sustaining and enhancing the capacity of social-ecological systems to cope with stress and adapt to uncertainty. For instance, approximately 10 million people today experience coastal flooding each year due to storm surges and landfall typhoons, and it is estimated that this number will reach to 50 million by 2080 because of the climate change and significant increase in population densities [Adger et al., 2005]. Therefore, the use of resources and maintaining sustainability status of systems are critical for the survival of future generations in all research fields.

Furthermore, disasters are threats to sustainability of communities and often to the environmental resources that those communities are depend on. In this context, a community with a global or local scale can be a continent or a country, an urban district or a co-operated environment. Any community is engaged with the policy management problems ranging from risk management, preparedness and responses to decision making strategies, data integration
Semantic heterogeneity is the identification and integration of semantically related information, which is represented in different resources. The problem is caused by various representations of the same or overlapping information. In other words, during the data integration process, semantic heterogeneity issues may arise at different levels such as terminology, structure and concepts of knowledge modelling, that are tackled in various disciplines such as databases, Web documents, knowledge systems and social media [Hull, 1997; Pisanelli et al., 2002; Magee, 2010].

Taking knowledge systems and social media as examples, the domains of sustainability and disaster management showcase the general problems of semantic heterogeneity in the aforementioned disciplines, which have gained a lot of attention in recent years. Various studies consider an overlap between the two domains with the focus on resilience and risk management, which highlights the severity of how humans adapt to a changing world. In recent years, topics such as maintaining the sustainability of communities, preparing for natural disasters, coping with emergency situations and responding to risk by decision making techniques have become critical. These have initiated a variety of research topics by the governments globally and locally.

Our motivation in this thesis is to examine one of these research topics on resolving semantic heterogeneity issues for two related domains of sustainability and disaster management by applying semantic technologies.

1.1 Motivation and Solutions

Knowledge representation requires interdisciplinary approaches to extract and reproduce important information from content. It aims to provide a suitable schema for extracted information that can be used for decision making tasks and constructing knowledge bases. It has been argued that a knowledge base reflects reusability, versatility and extensibility of a specific domain knowledge [Kumazawa et al., 2009]. Various methods have been proposed to construct knowledge bases [?Angele et al., 1998; Eriksson et al., 1995].

One of the methods to construct knowledge bases is the use of ontologies. In this thesis,
an ontology is referred as a specification of a conceptualisation (as defined by Gruber [1995]). Ontologies provide a structure of concepts, relations and instances for the domain knowledge, which are often developed through a set of ontology engineering steps [Grüninger and Fox, 1995; Staab et al., 2001b; Hovy, 2002; Gómez-Pérez et al., 2004]. In addition, complex design problems can be addressed by the use of ontology design patterns [Gangemi and Presutti, 2009]. Furthermore, an ontology must be evaluated against a set of evaluation criteria [Guarino and Welty, 2002; Lozano-Tello and Gómez-Pérez, 2004; Yu et al., 2009].

Given the role of ontologies in developing knowledge bases, a major challenge here is that researchers still question the structure of knowledge. A traditional definition of knowledge in Artificial Intelligence (AI) presented by Newell [1981] is that it is the information assigned to a system or an agent. The agent, here, is capable of computing its behaviour based on the principles of rationality. This perspective was challenged in early 1990s when Clancey [1990] pointed out that the primary aim of knowledge engineering should be modelling of different views of the world rather than reflecting one single expert mind. This perspective later appears as semantically describing the same real world concept from various domains, from which emerges the idea of integration of multiple heterogeneous resources.

The issue of semantic heterogeneity in the domains of sustainability indicators and disaster management is under-explored in this thesis. We present a model, that suggests computational solutions to the common issues of heterogeneity in the two related domains of sustainability indicator sets and disaster management. As displayed in Figure 1.1, our model proposes sustainability as the preparedness to risks and natural hazards, whereas disaster management is considered as the response to these situations. We also include communities, which confront risks, in the picture.

Two examples of communities used in this research are users of Circles of Sustainability and OzCrisisTracker applications, in which, this research is a part of a linkage project on Accounting for Sustainability: Developing an Integrated Approach for Sustainability Assessments, supported by the Australian Research Council (ARC) which aims to develop a “semantic” approach using open source software and systems, for supporting sustainability
Research Questions

reporting within organisations\(^1\). There are six different partners associated in this project: Angusta\(^2\), Cambridge\(^3\), City of Melbourne\(^4\), Microsoft\(^5\), Fuji Xerox\(^6\) and Cities Program\(^7\). These and other organisations can use sustainability indicators to measure the impacts of the projects and practices in economic, environmental and social terms. Such measures can form the basis for improving resource usage and efficiency, adapting policy, and institutional behaviour change.

In addition, a part of this research is supported by the IBM Research – Australia and funded by the Australian Computer Society (ACS) in developing a geotagger for OzCrisis-Tracker (that is an adaptation of CrisisTracker [Rogstadius et al., 2013]) that has extracted situation awareness reports from tweets collected during the Australian Bushfire and Flood seasons 2013-2014. These reports can be used to improve response times and effectiveness by emergency services.

1.2 Research Questions

Using the above model shown in Figure 1.1, we identify several semantic heterogeneity issues with regards to the data modelling and representation of the two domains, which lead us to define three specific research questions as presented below:

RQ1: How to develop an ontology for constructing a knowledge base that systematically represents semantic heterogeneity of sustainability indicator sets?

RQ2: Assuming to have found a positive answer to RQ1, what mechanism can be used in a sustainability reporting framework to help end-users for navigating and browsing heterogeneous sustainability indicator sets?

RQ3: How to use knowledge bases to automatically geotag emergency tweets which contain semantic heterogeneity in their location?

\(^3\)http://cambridgecollege.com.au/
\(^7\)http://citiesprogramme.com/
The overview of these research questions are illustrated in Figure 1.2, that reflects how real users can benefit from the outcome of this research. Question 1 tackles the heterogeneity issue of knowledge representation of sustainability indicator sets by using ontologies to construct a domain knowledge base. In building the knowledge base, a series of ontology engineering, design patterns and evaluation methods are used, resulting in the development of two ontology design candidates: specific and generic. Question 2 uses the generic ontology design to resolve the heterogeneity issue of user interface representation of sustainability indicator sets by developing software for a reporting framework (Circles of Sustainability given in Section 3.2.1). For browsing heterogeneous sustainability indicators, two interface mechanisms are presented: Tabular and Circular views. The comparative usability of these views are evaluated through a user study experiment. Question 3 resolves heterogeneity issues in georeferences of tweets’ content in the context of disaster management for an application (OzCrisisTracker given in Section 3.3.1) by the use of new and existing knowledge bases and ontologies.
1.3 Research Contributions

The overall contribution of the thesis is the use of semantic technologies to resolve representation issues of heterogeneous resources for two related domains of sustainability and disaster management. More specifically, the following key contributions are made.

1. This thesis guides the construction of two ontology design candidates (generic and specific) for the domain of sustainability indicator sets by applying an adopted ontology engineering approach and existing ontology design patterns. It shows, where an ontology needs to be designed for both seen and unseen indicator systems, a generic and reusable design is preferable. In addition, the thesis establishes how the use of existing structured approaches like METHONTOLOGY [Gómez-Pérez et al., 2004] and ROMEO [Yu et al., 2009] can reduce ambiguity in ontology design and evaluation for domain-level ontologies.

2. The thesis establishes how the generic ontology design can be used in developing two interface mechanisms (Tabular and Circular views) for browsing heterogeneous sustainability indicator sets in a reporting framework (such as Circles of Sustainability). Through a user study experiment, it shows that the Circular view can better assist expert and less-expert users to browse heterogeneous sustainability indicators sets.

3. The thesis shows how existing geographical knowledge bases can be used to develop a geotagger that automatically detects and resolves the heterogeneity of georeferencing in tweets’ content with high coverage and fine granularity. It also introduces a semantic annotation model that can improve the disambiguation of such heterogeneity issues in the context of disaster management.

1.4 Thesis Overview

The rest of the thesis is organised as follows.

Chapter 2 reviews and compares related work in various fields including: semantic heterogeneity, constructing knowledge bases, ontologies, ontology engineering methodologies,
ontology design patterns, ontology evaluation approaches and in the domains of sustainability and social media in the context of disaster management, which lead us to address the research questions.

Chapter 3 elaborates the three research questions together with the analysis of steps and methods that are used to address them.

Chapter 4 addresses the first research question by developing a knowledge base for sustainability indicator sets utilising ontologies. It steps through our proposed adaptable ontology engineering approach of METHONTOLOGY that results in two ontology design candidates: generic and specific. We then evaluate both ontology design candidates using the ROMEO approach.

Chapter 5 develops and evaluates the Circles of Sustainability software and its user interface by applying the generic ontology design to address the second research question, that is, which mechanism can help end-users to browse heterogeneous sustainability indicator sets. We first develop the user interface by providing two methods of browsing indicator sets: Tabular and Circular views. We then conduct a user-study for evaluating the overall usefulness of the software and ease of use of the two ways of browsing indicator sets.

Chapter 6 addresses the third research question, which is resolving heterogeneity of georeferencing tweets’ content in the context of disaster management. We present the OzCT geotagger that automatically detects the location(s) mentioned in the content of tweets with three possibilities: definite, ambiguous and no-location. Our method also semantically annotates the tweet components utilising existing and new ontologies.

Chapter 7 concludes the research findings and presents some future directions.
Chapter 2

Knowledge Bases and Ontologies

Science is a way of thinking much more than it is a body of knowledge.

—Carl Sagan

The overall goal of the thesis is to attempt to resolve the semantic heterogeneity issues that appear at the data and user interface levels for two related domains of sustainability and disaster management. In particular, our focus in this research is operationalising sustainability indicator sets and geotagging emergency tweets by developing knowledge bases that utilise new and existing ontologies. Accordingly, three research questions are presented in the previous chapter. This chapter provides a background of related work in different fields that lead us to address the research questions.

2.1 Semantic Heterogeneity

Integration and reconciliation of data represented by heterogeneous resources is a pressing topic in different domains of science. It can be described as a problem of semantic heterogeneity, referring to various representations of the same or overlapping data. The problem first gains attention in relational databases when data duplicated across multiple resources needs to be integrated. Kashyap and Sheth [1997] point out the problem of semantic heterogeneity as “identification of semantically related objects in different databases and the resolution of schematic differences among them”. Modern database management systems (DBMSs)
[Ramakrishnan and Gehrke, 2003] are unable to solve semantic heterogeneity problems in databases that are distinguished in two categories [Hull, 1997]: the appearance and logical structures. From the perspective of hardware and software platforms of a database, the issues are seen at the Application Programming Interfaces (APIs) level, which the solutions are sought in network communication protocols [Date and Darwen, 1987]. From the view of logical structure of a database, such issues appear at the data and schema level, therefore various conceptualisations and database schemas [Carey et al., 1995; Genesereth et al., 1997; Levy et al., 1996] are proposed in response. The second type of semantic heterogeneity issues are seen in Web documents, which require the integration of structured and semi-structured data resources [Bergamaschi et al., 2001]. In this form, semantic representation of data is not only an issue, but also the variation of data structures represented by different resources adds to the complexity. Some solutions are suggested in the development of methods, techniques and languages for Web content [Buneman, 1997; Buneman et al., 1996; Calvanese et al., 1998].

More recently, the proliferation of data systems on the Web has causes bigger challenges for the integration of knowledge systems. For instance, Pisanelli et al. [2002] discuss issues with regards to the integration and sharing of an extensive amount of information stored in various system repositories, in which each system has its own structure. Data integration in this context is related to interoperability, which is raised when the collaboration between systems and groups increase. One solution is suggested for data sharing through standardised communication between people and machines. A good example is the communication and exchange of information, which is a key requirement in business applications [Bergamaschi et al., 2001].

In a parallel argument, Magee [2010] raises the concern of assessing the commensurability of knowledge systems by developing an implied theoretical approach accompanied by a heuristic and analytic framework. Magee [2010] suggests using the concept of ontological culture in constructing formal knowledge systems by their communities (cultures) to represent the vague kinds of entities used in knowledge systems. Additionally, they suggest that the development of interlanguages for translation between the micro-languages within knowledge
systems, could reduce their exponential complexity.

Furthermore, understanding the data heterogeneity is also a major issue in analysing social media content. In this domain, heterogeneity defines as an inherent problem in the geo-scientific areas because of the wide variety of potential applications. Although significant improvement on this topic has been made [Cheng et al., 2010; Mahmud et al., 2012; Lee et al., 2013; Kiryakov et al., 2004], semantic issues are still insufficiently addressed. This issue is referred to as the need for *semantic interoperability among autonomous and heterogeneous systems* [Goh et al., 1999]. Two types of semantic heterogeneity for analysing social media are distinguished by Klien et al. [2006]:

- **Cognitive heterogeneity**: Due to various perspectives of the same real world facts there may not be a common base of definitions of the underlying facts between two disciplines (domains).

- **Naming heterogeneity**: The same real facts are understood in the same way but are named differently.

In the first type of the problems, realising the insights of specific situations requires a platform that can ingest multiple sources of data, analyse and correlate them using a variety of models and tools. In the second type, no shared vocabularies exist that specifically allow describing a topic of interest broadcasting in social media.

In other words, cognitive heterogeneity requires a social activity for reconciliation of underlying theories to ingest multiple sources and realise the insights of specific situations, whereas, naming heterogeneity can be partially solved by computational activities such as ontology matching techniques and translation algorithms that allow describing a topic of interest among heterogeneous resources.
2.2 A History of Development of Knowledge Base Systems

Development of knowledge-based systems (KBSs) starts in the early days of Artificial Intelligence (AI) [Studer et al., 1998] and continues until the present [Hendler, 2001; Berners-Lee et al., 2001; Shadbolt et al., 2006]. Reviewing literature highlights two major processes for constructing KBSs: the transferring process and the modelling process that are described next. There are many other approaches used in knowledge engineering community, such as, VITAL [Shadbolt et al., 1993], Commet [Steels, 1993] and EXPECT [Gil and Paris, 1994]. However we do not discuss them in this thesis because they are not widely used in the literature and are specific models for particular knowledge domains.

2.2.1 Transferring Process

In the early 1980s the development of a KBS was considered as a transferring process from human knowledge into a knowledge base. The main assumption in this process was that the knowledge already exists and just needed to be collected and implemented through various approaches such as interviewing experts [Studer et al., 1998]. However, this view failed to account for the combination of various knowledge types, and the lack of adequate justification of different rules. As a result, maintenance of such knowledge bases were difficult and time-consuming. Therefore, the transfer approach was only feasible for small prototypical systems, and, instead the modelling approach was introduced to produce large, reliable and maintainable knowledge bases.

2.2.2 Modelling Process

Current knowledge engineering approaches [Schreiber et al., 1993; Gaines et al., 1993] focus on the modelling process for knowledge engineering tasks. Unlike the transfer process, the knowledge in the modelling process is assumed to be not directly accessible but has to be built up and structured during the knowledge acquisition phase [Morik, 1991]. Guarino [1995] states that within the modelling process there is a correlation between a knowledge base and two separate subsystems: the agent’s behaviour (for example, the problem-solving expertise) and its own environment (the problem domain). Similarly, Studer et al. [1998] resemble
the modelling process of KBS to the development of a computer model with the purpose of realising problem-solving capabilities. Studer et al. [1998] outline several consequences of using the modelling process for knowledge structuring which are listed as follows.

1. The product of knowledge engineering should partially describe an inherent structure in the domain by taking other relevant domains into account.

2. The task-independent activity of domain analysis also results in facilitating interaction and communication between diverse agents.

3. Since knowledge engineering is an expensive procedure, data sharing efforts in modelling enhances the quality of formalized bodies of knowledge to be shared and reused for a variety of different purposes.

4. It is a cyclic process with the extension of refinement, modification or completion, and further acquisition of knowledge may be required.

5. Since the modelling process is dependent on the subjective interpretations of the knowledge engineer, the procedure is sometimes faulty and an evaluation method with the possibility of revising the model in every stage is essential to generate an adequate model.

Three frameworks using the modelling process presented by Studer et al. [1998] are: CommonKADs [Schreiber et al., 1994], MIKE [Angele et al., 1998] and PROTÉGÉ-II [Eriksson et al., 1995]. In brief, CommonKADs [Schreiber et al., 1994] is a developed version of KADs [Schreiber et al., 1993] that is a collection of models – Organisation, Task, Agent, Communication, Expertise, and Design – where each model represents a specific aspect of the KBS and its environment. The Model-based and Incremental Knowledge Engineering (MIKE) approach [Angele et al., 1998] suggests a development procedure for KBSs that integrates semi-formal and formal specification techniques and prototypes them into an engineering framework. Studer et al. [1998] outline the main difference between the two systems, where MIKE supports an incremental reversible system development; whereas, CommonKADs suggests a model-based framework. The third approach, PROTÉGÉ-II [Eriksson et al., 1995],
contains a task-method-decomposition structure, that provides the KBS's development with reusing of Problem Solving Methods (PSMs) [Birmingham and Klinker, 1993] and ontologies. It applies a PSM on a task to decompose it into corresponding subtasks, which can be solved by primitive methods. Two types of ontologies are used in PROTÉGÉ-II. The first type is method ontology and can be used as input and output of a method, which specifies concepts and relations reused by a PSM. The second type of ontology used in PROTÉGÉ-II is domain ontology, that defines a shared conceptualisation of a domain.

2.3 Ontologies

The term ontology has recently received the popularity within the knowledge engineering community, although its definition still remains ambiguous as it is used in different disciplines from various perspectives over time. Giaretta [1995] summarises seven definitions proposed for an ontology.

1. Ontology as a philosophical discipline.

2. Ontology as an informal conceptual system.

3. Ontology as a formal semantic account.

4. Ontology as a specification of conceptualisation.

5. Ontology as a representation of a conceptual system via a logical theory
   
   (a) characterised by specific formal properties.
   
   (b) characterised only by its specific purposes.

6. Ontology as the vocabulary used by a logical theory

7. Ontology as a (meta-level) specification of a logical theory.

The interpretation of 1 refers the origin of the term ontology (or ontologia) which is back to a branch of Philosophy. It was used by early students of Aristotle who referred to an ontology as a synonym of "metaphysics". The study discusses what might exist and seeks
to the essence of things through changes. In this context, ontology provides a definitive and exhaustive classification of entities in all kinds of being. The interpretation of 2 sees an ontology as a casual representation of a conceptual system which cannot be used for other systems due to informality, whereas, the interpretation of 3 makes the representation formal at the semantic level for a specific system. However, such an ontology is still useless for other systems because it can only be used for a system with the similar structure of the target system. Some examples of formal semantic ontologies are presented in these studies [Guarino et al., 1994; Guarino, 1995]. Interpretation of 4 sees an ontology as a description of concepts and relations that are exist for an agent or a community of agents, which is mostly reflected in AI from computing discipline. Definition 5 looks at an ontology as a logical theory whether it contains particular formal properties or it presents the new changes. Definition 6 sees an ontology not only as a logical theory but also as a set of vocabularies that is used by a logical theory. This perspective is similar to the view 5.1 in which an ontology is thought of a specific collection of vocabularies consisting a set of logical definitions. In that sense, definition 4 is similar to the view 5.2, seeing conceptualisation as a list of particular vocabularies with built-in relations. Finally, definition 7 sees an ontology as a meta-level specification of a logical theory, which specifies the “architectural components” within a particular domain theory.

2.3.1 Ontology Components

Given the definition of an ontology as a specific conceptualisation in a shared domain, one of the fundamental aspects of ontology is presenting higher level distinction of concepts which assists to understand the lower level concepts of a domain knowledge. Ontology components are proposed during the ontology development to pursue this goal. Fernandez et al. [1997] defines ontology components as: concepts, relations, instances, attributes, constants and formal axioms. The first three are essential elements to construct an ontology, whereas, the other three components are optional depending on the domain and application levels.

**Concepts** (or classes) represent key elements of a specific domain. In other words, a concept is an entity with a key role that can be inherited by other entities. Classes can be
abstract or specific. For example, in sustainability domain, \texttt{IndicatorSet} is an abstract concept that can be also looked at as a specific concept.

**Relations** specify the types of association between concepts of the domain. In ontologies, relations are binary, in which the first argument is known as the \textit{domain} of the relation and the second argument refers as its \textit{range}. Binary relations can be applied for different types of associations including: (i) relations that are used for building class taxonomy such as \texttt{subclass-Of} and \texttt{is-a}. To be more specific, \texttt{is-a} relation holds only between classes. \texttt{< class A is-a class B >} holds if the instance set of A is a subset of the instance set of B. (ii) relations that connect two different classes. For instance, in sustainability domain, the \texttt{Indicator} class is linked to the \texttt{IndicatorSet} by the relation of \texttt{belongsToIndicatorSet}.

**Instances** (or individuals) are the elements in an ontology that are instantiated from concepts. Mizoguchi [2004] defines the relation between a class and an instance, as a conceptualization of a set element of \textit{x} of \textit{X}, if and only if each “intrinsic property” of \textit{x} satisfies the intensional condition of \textit{X}. And, then and only then, \texttt{< x instance-of X >} holds.

**Attributes** are those relations that are used as concept properties which are also known as \textit{slots}. The range of relations is a concept, whereas, the range of the attributes are datatypes such as \texttt{string}, \texttt{numbers}. An example of an attribute in sustainability domain is \texttt{UnitOfMeasurement}, which its datatype is a \texttt{string}.

**Constants** refer to those attributes or instances that always have the same value. For instance, a constant can be an entity that is invariant across ontology instances.

**Formal Axioms** represent the conditions between relations and concepts that are always true in the knowledge base. Axioms – also referred as entities or objects – appear at the \textit{symbol level} of ontology development, which are required to be formulated in first-logic languages or be modelled for some cases of bigger ontologies.

2.3.2 Taxonomies

\textit{Taxonomies} refer as the centre of most conceptual models. According to Welty and Guarino [2001], well-structured taxonomies are the key elements in reusing and integrating models and
Ontologies are particularly useful for human interpretation, whereas improperly structured taxonomies results in models which are confusing and difficult to reuse or integrate. In ontology development, a taxonomy is referred to a hierarchical structure of a set of concepts and subsumption relations between them. These concepts are then represented as classes in the ontology. Broad concepts compose superclasses, that are inherited by subclasses using is-a relation. Unlike ontology, in taxonomy there is no distinction between the conceptual entities whether for instance, they ought to be represented by classes or individuals, or related to generic attributes or specific relations. In other words, the additional relations between concepts are not represented in a taxonomy. In Section 2.9.1, we provide some examples of taxonomies used in the sustainability domain.

2.3.3 Ontology Specification Languages

An ontology specification language explicitly express ontologies that are computationally understandable and decidable. Early ontology languages were KIF\(^1\), LOOM\(^2\) and F-Logic [Kifer et al., 1995]. Ontology languages are then developed for the Web content such as SHOE\(^3\), XOL\(^4\), DAML\(^5\) and OIL\(^6\). They, however, have some limitations: First, they are proposed for specific user groups and cannot be always easily extensible. Second, each language has its own specific syntax, that is difficult to integrate with other languages. Finally, their syntaxes are also varied in expressiveness and computational properties, therefore, these languages are complex to compute. Given these limitations, better alternatives for an ontology language in Semantic Web society are then suggested.

Web Ontology Language

Web Ontology Language (OWL) is introduced by by the World Wide Web Consortium (W3C)\(^7\) in 2004 to process machine interpretability of the Web content on the basis of

---

1 http://www-ksl.stanford.edu/knowledge-sharing/kif/
2 http://www.isi.edu/isd/LOOM/LOOM-HOME.html
3 http://www.cs.umd.edu/projects/plus/SHOE/
4 http://www.ai.sri.com/pkarp/xol
5 http://www.daml.org
6 http://www.ontoknowledge.org/oil
7 http://www.w3.org/TR/owl-features/
the previous languages, such as DAML, OIL and DAML+OIL which makes OWL to be compatible with Description Logics (DLs)\textsuperscript{8}. As a result, OWL documents can be reasoned by different DLs reasoners. More importantly, it supports existing Web standards like Extensible Markup Language (XML)\textsuperscript{9}, XML Schema\textsuperscript{10}, Resource Description Framework (RDF) and RDF Schema (RDF-S)\textsuperscript{11}. What makes OWL a stronger language than other ontology specification languages, is its ability to reference a very large number of Web resources. This integration facility of OWL with the Web standards makes it a more stronger language than other ontology specification languages.

OWL is substituted into three expressive sub-languages: OWL Lite, OWL DL and OWL Full. These versions of OWL are designed for the use of specific communities. OWL Lite provides classification hierarchy and simple constraint features. OWL DL provides maximum expressiveness without losing computational completeness. In other words, in OWL DL all computations will finish in finite time. One of the constraints of OWL DL is the type separation, in which a class can not also be an instance or a property, and a property can not also be an instance or class. OWL Full, on the other hand, supports users who want maximum expressiveness and the syntactic freedom of RDF with no computational guarantees. This feature makes this type more flexible that one class can be treated as simultaneously as a set of instances, and as an instance in its own right.

Resource Description Framework

Resource Description Framework (RDF)\textsuperscript{12} is a standard model of semantic Web developed by the World Wide Web Consortium (W3C)\textsuperscript{13}. RDF extends the structure of the Web resources with the use of a Uniform Resource Identifier (URI) to semantically link them in a triple format of “Subject”, “Object” and “Predicate”. Using this simple model, it allows structured and semi-structured data to be mixed, exposed, and shared across different applications. This

\begin{itemize}
  \item \textsuperscript{8}http://dl.kr.org/
  \item \textsuperscript{9}http://www.w3.org/XML/
  \item \textsuperscript{10}www.w3.org/XML/Schema
  \item \textsuperscript{11}http://www.w3.org/TR/rdf-schema/
  \item \textsuperscript{12}http://www.w3.org/RDF/
  \item \textsuperscript{13}http://www.w3.org/
linking structure also forms a directed, labeled graph, where the edges represent the names and, the links between two resources, represented by the graph nodes. This graph view is the easiest possible mental model for RDF and is often used in visual explanations.

The next topic to discuss is ontology granularity in which various ontologies may define a set of concepts with higher abstraction or granularity than others.

2.4 Ontology Granularity

While the definition of general levels of granularity is always not consistent in ontology literature, three different structures can be distinguished: (i) Gandon [2001] presents a three-layer model of abstraction: a top layer consists of an ontology with abstract concepts, a middle layer defining concepts for an application and a bottom layer including concepts relevant to a specific task. (ii) Fonseca et al. [2002] outline four levels of ontologies: top-level ontologies describing general concepts, domain ontologies presenting domain-related concepts, task ontologies describing task or activity concepts, and at the bottom, application ontologies combine domain and task concepts. Alexakos et al. [2006] presents a distributed hypermedia of three layers: an upper search ontology layer consisting of basic concepts, a domain description ontology layer, and a semantic metadata layer dealing with heterogeneous hypermedia server data. Incorporating these ideas, in this research we use a three-layer of ontology granularity including upper, domain and application ontologies. A description of each layer are as follows.
2.4.1 Upper Ontologies

The necessity of a large, comprehensive and formal ontology in various fields of research leads to creating standard descriptions and terminology for the entities and events that describes the broader knowledge (e.g. world). These ontologies are used as umbrella structures organising concepts in lower levels of an ontology. Two ontologies that are developed specifically with the purpose of using as upper ontologies are SUMO [Niles and Pease, 2001] and DOLCE [Guarino et al., 2003].

The Suggested Upper Merged Ontology (SUMO) [Niles and Pease, 2001] is an extension of the Standard Upper Ontology (SUO) email list, which was a starter document by the IEEE standard Upper Ontology Working Group compromising researchers from computer science, artificial intelligence, philosophy and linguistics. SUMO provides definitions of general terms by merging existing ontological content into a single, comprehensive and cohesive structure. It aims at “promoting data interoperability, information search and retrieval, automated inferencing and natural language processing”.

A Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE) [Guarino et al., 2003] is an upper level ontology developed by the WonderWeb project team\footnote{wonderweb.semanticweb.org/} compromising a large number of European research groups. DOLCE aims at “capturing ontological categories underlying natural language and human common-sense” by providing a common reference framework to facilitate information sharing.

2.4.2 Domain Ontologies

Domain ontologies are general types of ontologies that are often developed in a wide range of disciplinary areas by aiming at reusability of a higher shared value. Domain ontologies are developed in various disciplines of Earth Science (SWEET) [Raskin and Pan, 2005], Agriculture and Food (Plant Ontology) [Avraham et al., 2008] and Bioinformatics and the Life Sciences being particularly prominent examples: SNOMED [Spackman et al., 1997], Gene Ontology [Ashburner et al., 2000] and Protein Ontology [Natale et al., 2011].
2.4.3 Application Ontologies

Application ontologies are very specific in nature and relate directly to details of a given application or task. Such ontologies combine domain and task concepts to meet the requirements of a particular application or framework. The examples of application ontologies are diverse in different disciplines. We review some of the domain and application ontologies that are used for the areas of study in this thesis in Sections 2.9 and 2.10.2.

In developing ontologies, ontology engineers often require guidance of how to construct and apply ontology components and appropriate ontology granularity. This brings us to the notion of ontology learning and ontology engineering methodologies, which are discussed next.

2.5 Ontology Learning Approaches

Applying semi-automatically or automatically learning algorithms to another resource for developing ontologies is referred as Ontology Learning. Maedche and Staab [2001a] were the pioneers of ontology learning using text documents named Layer Cake. In this method, ontology components are defined as layers which from bottom to up are: terms, synonyms, concepts, relations and rules. The same authors [Maedche and Staab, 2001b] later introduced another ontology learning approach for the domain of Semantic Web, which consists of the following steps: (i) Importing and reusing existing ontologies, (ii) Extracting major parts for the target ontology from other resources such as Web documents, (iii) Pruning the target ontology from existing concepts to be fitted for its primary purpose, (iv) Performing ontology refinement to complete the target ontology at the fine granularity, and (v) Applying the target ontology to the specific application for the design validation. In addition to the above approaches, there are ontology learning tools that are used to develop and maintain ontologies from structured and semi-structured data. Some examples are OntoLearn\(^{15}\), Text2Onto\(^{16}\) and OntoBuilder\(^{17}\).

\(^{15}\)https://code.google.com/p/ontolearn/
\(^{16}\)https://code.google.com/p/text2onto/
\(^{17}\)http://ontobuilder.bitbucket.org/
Now, we discuss approaches in ontology engineering, which is the steps of developing an ontology.

2.6 Ontology Engineering Methodologies

*Ontology engineering* refers to a set of principles that are related to the development of an ontology in a specific domain. An ontology engineering methodology describes a process of constructing ontologies including requirements gathering, development, refinement, repository management and evaluation.

2.6.1 Skeletal Methodology

*Skeletal Methodology* is a generic model for ontology engineering suggested by Uschold and King [1995], that is used as a foundation in most ontology engineering methodologies. It describes ontology development steps as: (i) Identifying purpose and scope, (ii) Constructing the ontology in three phases of capturing, coding and integrating existing ontologies, (iii) Evaluation and (iv) Documentation.

2.6.2 Formal Method for Ontology Engineering

Formal method for ontology engineering is presented by Grüninger and Fox [1995] which was used for building an enterprise model. Its process includes five phases that are described below.

1. **Apply Motivation Scenarios**: The requirements are gathered from analysis of the application scenarios.

2. **Define Competency Questions**: A set of competency questions from the given motivation scenarios are identified.

3. **Formalise Ontology**: The first order logic of the ontology and the related axioms are specified.

4. **Apply Competency Questions**: Competency questions for the evaluation purposes are applied.
5. **Verify Ontology Completeness:** It is determined whether the ontology satisfies a given competency question using a set of instances as the ground literals.

Although formal method encourages the use of formal logic, there are, however, several limitations with this approach outlined by Yu et al. [2009] including:

1. Expertises are required with an ontology specification language for formal logic expressions.
2. This method is depended on the existence of a given set of instances for addressing the competency questions.
3. This method is also limited to the use of an evaluation method with a formal logic language only.

Despite these limitations, formal method provides a starting point of ontology evaluation through using a set of competency questions which is followed in other ontology engineering approach that are described next.

### 2.6.3 On-To-Knowledge Methodology

On-To-Knowledge is presented by Staab et al. [2001b] as an on ontology development for knowledge management applications. The methodology includes five steps as follows.

1. **Feasibility Study** identifies foreseeable problem areas and possible solutions prior to the ontology development process are identified.
2. **Kickoff** describes ontology requirements including goals, scope, users, use cases, competency questions and possible existing ontologies.
3. **Refinement** develops the ontology according to the specification document.
4. **Evaluation** tests the ontology against competency questions until the ontology reaches a level of satisfaction.
5. **Maintenance** reflects changes and accommodating updates after the ontology has been in use for a period of time.
An advantage of this approach compared with the Skeletal methodology [Uschold and King, 1995] is that the documentation step is applied at the early stage of ontology development rather than near the end as the Skeletal methodology suggests. It is also used as the reference throughout the ontology engineering lifecycle. Furthermore, another benefit of this aspect is its cyclical process that allows reflecting changes and accommodating updates until a level of maturity is reached. The major limitation of this approach is that the ontology engineer decides on the metrics or experiments to be used to evaluate the ontology. This is problematic because it may vary from one application to another. Hence, there is no systematic or standardized method for determining metrics for evaluating and comparing ontologies.

2.6.4 SENSUS-based Ontology Methodology

Hovy [2002] develops an ontology engineering methodology based on the SENSUS ontology which excludes irrelevant concepts outside of a given domain and enables the output ontology to be extended and modelled additional domain knowledge. The steps are:

1. Identify seed terms.
2. Manually Link seed terms to SENSUS ontology.
3. Add paths and roots to the ontology.
4. Add new domain terms to the ontology that are not already presented.
5. Add the complete subtree of new nodes for the extended model.

The advantage of this ontology is grounding domain ontologies with the SENSUS ontology. This method, however, is not considered a complete procedure, as it does not discuss any steps for evaluation or documentation.

2.6.5 METHONTOLOGY

METHONTOLOGY is presented by Gómez-Pérez et al. [2004] for ontology development in the domain of chemicals. This method include three major steps as shown Figure 2.2:
Ontology Engineering Methodologies

**Knowledge Acquisition**
- Performing interviews with experts and text analysis to capture the key information of the domain knowledge.

**Integration**
- Comparing and integrating existing ontologies used in similar domains with the GT designed in the conceptualization activity.

**Evaluation**
- Evaluating the ontology using varying evaluation approaches and defining specific criteria to verify the “correctness” of the ontology.

**Documentation**
- Keeping record of a summary of design principles, challenges and refinement after completion of each development activity.

**Configuration Management**
- Relating to management activities that consists of arranging different settings for better control of the resources.

**Figure 2.2: METHONTOLOGY [Gómez-Pérez et al., 2004] Overview Adapted from [Corcho et al., 2005]**
1. **Management activities** are conducted at the start of the ontology development process, to identify the tasks to be performed, and the time and the resources required for their completion. Examples in this phase include control and quality assurance activities.

2. **Development activities** shape the basis of the ontology. Ideally these are conducted through small incremental and iterative cycles. Under the detection of any changes or mistakes, it is possible to return to any previous activities and make modifications and refinements. The development activities are: Specification, Conceptualization, Formalization, Implementation and Maintenance.

3. **Support activities** are carried out simultaneously with the development process activities. These activities include Knowledge Acquisition, Evaluation, Documentation and Configuration.

We have chosen this method to develop the ontology for sustainability indicator sets in this thesis. Our rational for this choice are described in Section 3.1.1.

### 2.7 Ontology Design Patterns

The primary concern in ontology engineering is building conceptual data schemata by capturing the key information of the domain knowledge. The aforementioned methodologies, however, cannot capture all the modelling problems within the domain knowledge. The notion of ontology design patterns is introduced to overcome these complexities.

Current ontology engineering methodologies [ Guarino and Welty, 2002; Lozano-Tello and Gómez-Pérez, 2004 ] lack to describe the generic concepts, that are the abstraction of their design. In software engineering, the abstraction is offered by *Software Design Patterns (SDPs)* [ Gamma et al., 1994 ]. Using SDPs in software design encapsulate the strengths of models with well-understood properties, which results in constructing abstract models that solve the generic design problems in *Object Oriented Programming* [ Booch, 1994 ]. Such designs with efficient engineering process then generate high-quality software artefacts. *Ontology design patterns (ODPs)* represents the same principle in ontology development. An ontology ideally
is a composition of different related ODPs, which are resembled as building blocks constructing the ontology structure. Recognising generic or abstract ontology components are the key point to specify appropriate ODPs. This process is domain-dependent and requires deep understanding of the key concepts of the domain problem. Similar to SDPs, ODPs are abstract solutions (also referred as small use cases) to known problems in the field of ontology engineering [Aranguren, 2005]. However, given that ontology engineering is a newer field to compared with software engineering, the definition, representation and application methods of ODPs lack the same level of agreement of software engineering design patterns.

The literature about ODPs can be divided into two areas: (i) studies that discuss the notion of ODPs and (ii) research that represents concrete ODPs for tackling specific design problems in developing ontologies.

Reich [2000] begins the discussion about the notion of ODPs in the field of molecular biology. Later, the idea of Semantic Patterns and Knowledge Patterns [Staab et al., 2001a] begins to be presented as reusable components for building knowledge bases. It is followed by a series of research by Gangemi [2005]; Gangemi et al. [2007] that distinguish between Logical, Conceptual and Content Ontology Design Patterns. Finally, Gangemi and Presutti [2009] revisit the patterns and classify them into six major categories:

1. **Structural ODPs**

   - **Logical ODPs** are content-independent and expressed in logical language such as OWL. However, they are affected by formal structure of the domain knowledge.

   - **Architectural ODPs** influence the overall shape of an ontology by suggesting design choices based on specific needs.

2. **Correspondence ODPs**

   - **Re-engineering ODPs** suggest design solutions to transform a conceptual model – which can even be a non-ontological resource – into a new ontology.

   - **Alignment ODPs** create semantic associations between two existing ontologies.
3. **Content ODPs (CPs)** provide design solutions to domain modelling problems by instantiation or compositions of Logical ODPs. However, unlike OPs which are content-independent, CPs are content dependent and can be reused as building blocks in other conceptual domains.

4. **Reasoning ODPs** are used as “applications” of Logical ODPs that apply the behaviour implemented in a reasoning engine to obtain certain results. Some examples of Reasoning ODPs are: classification, subsumption, inheritance and materialization.

5. **Presentation ODPs**
   - **Naming ODPs** provide “homogeneity” in naming classes and relations to enhance the ontology readability by humans. For instance, namespace creation from a URI such as http://www.fao.org as FAO, is a type of Naming ODPs.
   - **Annotation ODPs** improve understanding of an ontology and its components by providing annotation properties. For example, the use of rdfs:label is a type of Annotation ODPs.

6. **Lexico-Syntactic ODPs** refer as linguistic structures which contain certain keywords in a specific order. These ODPs are useful in combining simple Logical and Content ODPs with natural language sentences by generalizing the facts.

With regards to the literature that presents concrete design of ODPs, most attempts are individual examples proposed by Semantic Web Best Practices and Development Working Group\(^\text{18}\). Over time, ontology engineers develop an ontology network, Semantic Web Portal\(^\text{19}\) which initially is developed for the NeOn project\(^\text{20}\). This forum is currently used between ontology developers to review and update existing ontology design patterns and submit new patterns for new ontology modelling problems in various domains.

Evaluating ontologies is an essential step in most ontology engineering methodologies, that is discussed next.

\(^{18}\)http://www.w3.org/2003/12/swa/swbpd-charter
\(^{19}\)http://ontologydesignpatterns.org/
\(^{20}\)http://www.neon-project.org/
2.8 Ontology Evaluation Methodologies

The primary concern of Ontology evaluation methodologies is to examine whether requirements for a given ontology have been met. Five main requirements in ontology design and ontology evaluation are presented by Gruber [1995]: i) Clarity means that an ontology should “communicate the intended meaning of the defined terms”, ii) Coherence means the axioms used in the ontology should be “consistent with the ontology definitions and concepts”, iii) Extensibility refers to an ontology with the possibility of using “a shared vocabulary”, iv) Minimal encoding bias means minimising the conceptualisation at the knowledge level without “depending on a particular symbol-level encoding”, and v) Minimal ontological commitment means that an ontology should make “a few claims as possible about the world being modelled, allowing other parties to instantiated if needed”.

In addition, there are other goals in ontology evaluation, such as ontology selection and tracing progress in ontology development. In this section, we present two ontology evaluation methods, OntoClean and OntoMetric, that have been influential in ontology evaluation techniques to date.

2.8.1 OntoClean

OntoClean is presented by Guarino and Welty [2002] which is domain independent aiming at simplifying modelling assumption to validate taxonomies. Applying OntoClean results in ontologies to meet requirement of correctness, that is whether the entities of the world being modelled are correctly represented in the ontology. OntoClean introduces a set of meta-properties such as rigidity, identity and unity, which capture various characteristics of classes to assess the correct usage of the subsumption relations determining whether they should be modelled as properties, as a subclass of another concept or even a separate class on its own [Guarino and Welty, 2002]. There are some limitations with OntoClean. One of the constraints is that each class in the ontology needs to be manually classified and tagged with the proposed meta-properties, which is a tedious and time-consuming activity. In addition, the classification task is subject to the ontology engineer’s interpretation of representing concepts in the ontology. Therefore the evaluation results may vary from one engineer to
another. Finally, OntoClean only evaluates the correctness of the subsumption relations of an ontology and it fails to validate other aspects of ontology such as completeness and soundness.

2.8.2 OntoMetric

OntoMetric is a criteria-based ontology evaluation methodology presented by Lozano-Tello and Gómez-Pérez [2004] that is used for ontology selection based on a set of characteristics. It differentiates ontologies using an adapted version of the Analytic Hierarchy Process (AHP) method that is a decision-making process based on multiple criteria [Saaty, 1990]. The OntoMetric process is a multi-level tree (or a taxonomy) of characteristics (MTC). The top level of this taxonomy contains five dimensions that are used to construct an ontology including: content, language, methodology, tools and the cost. A set of factors are associated with each dimension in the middle level of the hierarchy. And a set of characteristics obtained from existing work are also associated with each factor at the bottom level. The steps of utilising OntoMetric are:

1. Analyse project aims.
2. Obtain a customised MTC.
3. Weigh up each characteristic against each other.
4. Assign linguistic score for each characteristic of a candidate ontology.
5. Select the most suitable ontology.

However, there are several limitations associated with OntoMetric. First, specifying the customized MTC for ontology selection is subjective to the user’s manual specification and as a result it is inconsistent. Second, the set of characteristics for evaluating content is limited. In addition, there is a lack of specific measures for scaling linguistic scores and it depends on the user to assign values for the characteristics of the candidate ontology. The third and perhaps the most important drawback of OntoMetric is that this method can be only used for decision-making tasks for the most suitable ontology from a set of candidates ontologies.
Finally, in the case of a new ontology being developed, OntoMetric no longer can be used. The latter limitation is our main reason not to select this methodology for the ontology evaluation method in this research because we plan to extend the presented model in the future, therefore we seek for an ontology evaluation technique that meet this extendibility requirement.

2.8.3 ROMEO

Requirements Oriented Methodology for Evaluating Ontologies ROMEO presented by Yu et al. [2009] is a tasked-based ontology evaluation method, that is used for evaluating OSIS.

As illustrated in Figure 2.3, the evaluation tasks of ROMEO starts by establishing a set of ontology roles to clarify the needs of a given ontology. According to Yu et al. [2009], eliciting the roles of an ontology is important to understand how the ontology is used in the context of an application and it also helps to determine a set of appropriate ontology requirements. Next, these roles are linked to a corresponding set of ontology requirements. Yu et al. [2009] specify a series of generic requirements, such as “competency”, “capability”, “functionality” and “standardized”. Ontology requirements must reflect a set of needs from a suitable ontology for the given application. These in turn are mapped to a set of criteria-related questions. The ROMEO approach stipulates that a set of questions is administered for each
of the requirements. Such questions explore various aspects of a given requirement providing a
deeper understanding of the ontology. In addition, criteria questions have lead to appropriate
measures which are critical in an ontology evaluation process. Yu et al. [2009] propose a list
of criteria questions for a variety of ontology requirements. Ultimately, measures are then
developed to answer these questions, and hence, whether a given requirement has been met.
Various ontology criteria measures have been proposed in the ontology evaluation literature
that are summarised in a previous survey conducted by Brank et al. [2005].

We use this approach to evaluate the ontology to develop for sustainability indicator sets.
The details are given in Section 3.1.5.

2.9 Ontologies used in Sustainability and Sustainability Reporting Indicator
Sets

Increased attention to the concepts of sustainability and sustainable development in different
studies [Meadows et al., 2004; Burton, 1987] has encouraged the emergence of sustainability
reporting. The term sustainability indicator, closely related to the concepts of sustainability
and sustainability indicators discussed in Chapter 1, refers to a general set of descriptive
reporting practices, from “top-down” annual reporting against standardised indicator sets,
to ad hoc, one-off or semi-periodic assessments against “bottom-up” and locally-grown mea-
sures [Fraser et al., 2006]. According to Magee et al. [2013] top-down approaches produce
generalised sustainability reports by global focus, while bottom-up approaches tend to be
adopted by non-government organisations (NGOs), sub-national or municipal authorities,
and community groups [Agger, 2010]. The similarity of both approaches is found in the
use of framework-based technique including a collection of measures, procedures, tools and
principles that guides reporting practices.

There have been several attempts to develop domain and application ontologies in the
context of sustainability and sustainability reporting indicator sets. Brilhante et al. [2006]
present an ontology for the domain of Indicators and Sustainable Development (ISDs)21 with
the emphasis on the economic dimension (ISD-Economics). This ontology is developed on

the basis of the UN report on countries undertaking sustainable development programmes [Kuik and Gilbert, 2002]. ISD-Economics addresses several IDSs system concerns such as: providing an appropriate vocabulary for the user interface with a hierarchical organisation of the domain, and operational performing on indicators using properties and relations defined in the ontology so on. Similarly, Madlberger et al. [2013] develop an ontology for the domain of Corporate Sustainability Information Systems. This ontology fulfils three requirements: (i) Providing an intuitive representation of Corporate Sustainability concepts, (ii) Reflecting the hierarchical structure of economic indicators, and (iii) Generic mapping of sustainability heterogeneous resources – such as GRI-XBRL taxonomy – into the predefined concepts of the ontology.

Several application ontologies are also developed for specific tasks. Kumazawa et al. [2009] outline an ontological approach to structure the concepts and relations within the field of sustainability science. Han and Stoffel [2011] suggest an ontology for the integration of qualitative case studies in the domain of environmental sustainability research. Similarly, an ontological approach is presented by Pinheiro et al. [2009] to link sustainability indicators. In addition, Pinheiro et al. [2009] present a social ontology by relating and structuring environmental, social and economic issues in a way that measures overall social sustainability.

A very recent study on urban indicators is presented by Fox [2014], which develops a generic and reusable ontology for the ISO37120 Global City Indicators with the use of existing foundation ontologies and generated trans-foundation axioms.

To a large extent, sustainability indicators are introduced to address issues of critical conditions in complex systems. In other words, indicators can help to provide solutions for such issues including three key objectives [Kuik and Gilbert, 2002]:

1. Raising awareness and understanding of a system.
2. Informing decision-making.
3. Measuring progress toward established goals.

In addition to the above objectives, sustainability indicators are particularly structured for the purpose of organisational governance, where they are offered developed as simple
taxonomies rather than fully fledged ontologies.

Figure 2.4: A Snapshot of Sustainability GRI Indicator Set - Not all categories are shown

2.9.1 Taxonomies of Sustainability Reporting Indicator Sets

Well-known standardised frameworks for sustainability reporting and indicator sets include corporate reporting against the Global Reporting Initiative (GRI) indicators and guidelines\(^{22}\), the Organization for Economic Co-operation and Development (OECD)\(^ {23}\) and the United Nations Statistics Division (UN Social Indicators)\(^ {24}\). These organisations represent sustainability indicators in hierarchical structures including categories and sub-categories. The snapshots of the GRI and OECD taxonomies are shown in Figures 2.4 and 2.5 compro-

\(^{22}\)http://www.globalreporting.org/
\(^{23}\)http://www.oecd.org/
\(^{24}\)http://www.un.org/esa
Ontologies used in Sustainability and Sustainability Reporting Indicator Sets

Figure 2.5: A Snapshot of OECD Sustainability Indicator Set - Not all categories are shown

mising categories and sub-categories (or aspects and themes) and indicators at the bottom of the hierarchy. More specifically, GRI sustainability indicator set is presented in eXtensible Business Reporting Language (XBRL)\(^{25}\). XBRL is an XML-based language introduced to exchange business information. It uses the XML notation such as XML schema, XLink\(^{26}\) and XPath\(^{27}\) to express the semantic connections required in business reporting. Only very limited research investigates how to transform such a taxonomy into an ontology. For example, Madlberger et al. [2013] first used GRI-XBRL taxonomy\(^{28}\) as the basis to automatically generate the concepts of the Corporate Sustainability ontology.

\(^{25}\)http://www2.xbrl.org//au/
\(^{26}\)http://www.w3.org/XML/Linking
\(^{27}\)http://www.w3.org/TR/xpath/
\(^{28}\)https://www.globalreporting.org/reporting/reporting-support/xbrl/Pages/default.aspx
Having discussed ontologies and sustainability, in response to the first two research questions, we now turn to discuss the use knowledge bases and ontologies in social media in the context of disaster management, which is a background study to address the third research question.

2.10 Knowledge bases and Ontologies used in Social Media Content Analysis

The rapid and pervasive use of social media and advances in multiple platforms (mobile and Web-based applications) have made accessing and sharing information ubiquitous. In particular, semantic heterogeneity issues in social media appear at different categories on the basis of three criteria: the type of the connection between users, the ways of sharing information and the user interactions with the media stream [Bontcheva and Rout, 2012]. The categories are: Interest-graph media such as Mircoblogging, Social Network Sites (SNS) such as Facebook, Professional Networking Services (PNS) such as Linkedin, and Content sharing discussion services such as Youtube and Vimeo.

The social nature of human interactions is one of the reasons for the widespread adoption of social media. Individuals become part of a virtual community and communicate remotely by voicing their opinions via social networks. Taking microblogging as an example, at the Twitters 7th birthday on 21st March 2013, it had 200 million active users creating over 400 million tweets each day

In some previous studies, the role of social media, such as Twitter, is considered as a backchannel for transmitting information that cannot be received through the traditional media [McCarthy and Boyd, 2005; Mills et al., 2009]. White [2010] discusses how Tweets effectively link documents and pictures as well as locations if using a mobile device, which can effectively communicate the severity and range of a disaster.

Recent work emphasises the real time advantage of Twitter messages during emergencies, that is discussed next.

A tweet refers to a short text message of up to 140 characters. The topics of tweets posted by Twitter users across the world are various and cover almost everything from daily routine to the important events such as sports and politics.

https://blog.twitter.com/2013/celebrating-twitter7
Knowledge bases and Ontologies used in Social Media Content Analysis

In particular disaster-related tweets posted by victims, volunteers, and relief agencies are often distinguished from other types of tweets, because they contain information about an emergency situation or provide data in response to such circumstances. Although Twitter provides a geotagging feature to all its users, users do not necessarily enable this facility and only a limited number of tweets are tagged with geographical coordinates. The result of a one-month analysis of Twitter feeds in 2012 conducted by Lee et al. [2013] indicates that “99.1% of the tweets are not geotagged”. Moreover, the geolocation information often indicates the location of the user rather than the event mentioned in the tweet. Consequently, Twitter’s geolocation feature cannot be reliably used as indicator for event location. Therefore, content analysis of the tweets is an important area for identifying geographical references.

In the domain of disaster management, systems like Ushahidi\textsuperscript{30}, TweetTracker [Kumar et al., 2011], ESA [Yin et al., 2012], Twitcident [Abel et al., 2012], and CrisisTracker [Rogstadius et al., 2013] have been developed to assist humanitarian agencies and disaster relief workers. Ikawa et al. [2013] identify some of the advantages of using such systems for disaster agencies: (i) Listing situational reports extracted from social Web content, (ii) Performing operational activities such as recruiting volunteers, (iii) Providing emergency contacts and making decisions about distribution channels.

Hughes and Palen [2009] investigate Twitter activity from August to September, 2008 around the Democratic National and the Republican National Conventions and Gustav, and Ike hurricanes. Through conducting a set of experiments and data comparisons, they show that Twitter messages posted during the emergency and mass convergence events were very effective in information dissemination and supporting brokerage. In addition, they indicate that hurricane and convention tweets include more URLs due to their specific topics as compared to the general tweet pool.

In addition, two natural hazards events in North America in 2009 are examined with the aim of enhancing the situational awareness. One is the study by Vieweg et al. [2010], that investigate the Twitter messages related to these disasters and extract emergency features such as geolocation, location-referencing and situational update features. It has also been

\textsuperscript{30}http://www.ushahidi.com/
shown how the use of information extraction (IE) techniques in the emergency domains can improve situational awareness. These works also suggest that characterising computer-mediated communication (CMC) – such as microblog posts – can be used to reconstruct the relations between individuals and communities for emergency responses and communications.

Another specific study is the work by Starbird and Palen [2010], which examine the role of retweets for information propagation in a mass emergency. It is shown that retweeted messages are more used by locals to pass on emergency-related information rather than other types of tweets. Therefore, they argue that focusing on retweets may reduce noise during data collection of real-time emergencies. It is also concluded that retweets are important resources to analyse how information is differently valued from two aspects of being abstract and specific from an individual’s perspective.

Kaigo [2012] studies the effect of Twitter for the city of Tsukuba in Ibaraki prefecture during the Great East Japan Earthquake of March 2011. He argues that Twitter provided vital information and knowledge for the citizens of Tsukuba as a backchannel, while many of the lifelines were not functioning. It however showed that Twitter has the potential to transmit false rumours very rapidly during disasters and concluded that further consideration on accountability of the tweets’ content is required.

In what follows, we review related work that use two approaches in analysing social media content: Probabilistic language models and Ontologies.

### 2.10.1 Using Probabilistic Language Models

In the context of content analysis of social media, *probabilistic language models* are generally constructed based on the *features* collected from particular online data resources.

Serdyukov et al. [2009] develop probabilistic language models based on the terms people use to describe photos uploaded to Flickr on the World map\(^{31}\). These models associated with Bayesian inference estimate the location for a photo. They also demonstrate how the use of GeoNames database\(^{32}\) affects the decision whether a user’s submitted-tag is geo-related.

Similarly, Cheng et al. [2010] present a probabilistic framework that predicts the *city-level

\(^{31}\)http://www.flickr.com/

\(^{32}\)http://www.geonames.org/
location of each Twitter user. Two major components assist in developing the model: (i) a classification that automatically identifies words with strong local geo-scope in the tweet content, and (ii) a neighbourhood smoothing model that increases the accuracy of a user’s location estimate. Another approach is presented by Mahmud et al. [2012] that predicts the home locations of Twitter users at different granularities (For example, city, state or time-zone). Their algorithm first builds a set of statistical and heuristic classifiers. The former classifiers use words, hashtags or place names of tweets, and the latter classifiers apply the frequency of place names or Foursquare\textsuperscript{33} check-ins. An ensemble of the classifiers is generated to improve the accuracy of the prediction system. Recent research by Lee et al. [2013] builds a language model by deriving fine-grained location information from unstructured short messages given from FourSquare’s public API\textsuperscript{34}. The messages are associated with geolocation and tips left by individuals who check-in to the particular venues. Their probabilistic model is then used for tweet location prediction by applying a three-step method of Filtering, Ranking and Validating. Finally, the Emergency Situation Awareness (ESA) system [Yin et al., 2012] extracts important information from tweets broadcast by users regarding witnessing a crisis. The ESA system uses the Twitter API to capture tweets from regions of interest within Australia and New Zealand during the February 2011 Christchurch earthquake. It applies pre-processing and post-processing techniques on user profiles and the tweet contents and displays the results on a suite of visualisation interfaces for incident exploration. One of the main contributions of this work is building a statistical classifier on the basis of Support Vector Machine (SVM) classifier [Schölkopf et al., 2000] to identify tweets that report damage to roads, bridges, airports, power supplies, telecommunication and infrastructure. The geotagging module of ESA focuses on identifying the user’s geolocation rather than the geolocation mentioned in the tweet content.

2.10.2 Using Ontologies

In addition to probabilistic models, some domain ontologies are specifically created to model different kinds of social media such as user profile, content tagging, linking and other common

\textsuperscript{33}https://foursquare.com/

\textsuperscript{34}https://foursquare.com/
user behaviour. Semantically Interlinked Online Communities (sioc)\textsuperscript{35} models social communities, such as blogs, wiki and online forums, that include the key concepts of forums, sites, posts, user accounts, user groups and tags. Similarly, the recent extensions of sioc (sioct) specifically models microblogs by introducing the new concept of sioct:MicroblogPost with the property of sioct:follows to represent follows/followee relations and also sioct:tag property on Twitter. In addition, GeoNames Ontology\textsuperscript{36} makes it possible to add geospatial semantic information to the Word Wide Web. All over 8.3 million geonames toponyms now have a unique URI with a corresponding RDF statement.

Some existing work uses ontologies for semantic annotation of social media particularly in Twitter. Kiryakov et al. [2004] define semantic annotation as creating dynamic relations between ontologies and unstructured or semi-structured documents in a bidirectional manner. At the technical level, annotating in text means “applying all mentions of concepts from the ontology (i.e. classes, instances, and relations) on meta-data referring to their URIs in the ontology” [Bontcheva and Cunningham, 2011]. According to Bontcheva and Rout [2012], ontology-based Recognition Entity technique includes two major phases: Entity Annotation and Entity Linking. Entity annotation (also known as candidate selection) identifies all mentions in the text of classes and instances from the source ontology. Entity linking (also known as reference disambiguation or entity resolution) uses knowledge from the ontology or other linking open data resources to determine the conceptual information from the text that results in choosing the correct URI.

Furthermore, constructing task-based application ontologies for social media are addressed by a few researchers. Iwanaga et al. [2011] build an evacuation centre ontology using the Twitter API to retrieve data which contains explicit tags of particular evacuation centers. For capturing the relevant tweets they use the SVM technique [Schölkopf et al., 2000] and also utilise existing ontologies Vcard\textsuperscript{37}, Geo\textsuperscript{38} and Time\textsuperscript{39}. Dilo and Zlatanova [2008] present a UML-based data model of emergency management from a database built before

\textsuperscript{35}http://sioc-project.org/ontology
\textsuperscript{36}http://www.geonames.org/ontology/documentation.html
\textsuperscript{37}http://www.w3.org/2006/vcard/ns
\textsuperscript{38}http://www.w3.org/2003/01/geo/wgs84pos
\textsuperscript{39}http://purl.org/NET/c4dm/timeline.owl
the disaster. In order to facilitate communications at the time of disaster, this model stores specific features in particular classes such as people’s complaint in “Complaint”, property and life loss in “Process” and event information in “EventObject”. This ontology differs from other conceptual models in the spatial and temporal information representation as attributes of class Complaint rather than separate classes which adds simplicity and facilitates implementation in a DBMS. Wenjun et al. [2009] build an emergency response organization ontology model that inherits the concepts from SUMO [Niles and Pease, 2001]. The presented module contains five basic modelling meta-languages of ontologies including: concepts, relations, functions, axioms and instances which can define an emergency response organisation ontology as five-tuple.

2.11 Discussion

In reviewing the related work, we first began by defining semantic heterogeneity issues in various domains of databases, knowledge systems of social science and social media. We then looked at the history of knowledge base systems including: Transfer and Modelling processes. The main difference between these is the correlation between the agent’s behaviour and its own environment, which provides different problem-solving capabilities. In developing the knowledge bases in this thesis, we choose the PROTÉGÉ-II modelling framework [Eriksson et al., 1995] due to the following reasons: First, the use of Problem Solving Methods (PSMs) in constructing knowledge bases make the knowledge developed by domain experts explicit and as a result, the strategical knowledge can be reused for representing similar data structure and building new applications within the same or similar domain(s) [Studer et al., 1998]. Therefore, given the two areas of study in this thesis, sustainability and disaster management, we have develop knowledge base systems that can be reused in various applications rather than constructing a model to represent specific knowledge of a particular domain. Second, since both PSMs and domain ontologies are reusable components for constructing a knowledge base, PROTÉGÉ-II suggests the notion of an application ontology to overcome the interaction problem between domain ontologies and PSMs with their associated method ontologies. An application ontology extends domain ontologies with PSM specific concepts
and relations [Gennari et al., 1994]. We find the further development of an application ontology in PROTÉGÉ-II approach, an advantageous aspect that adds extensibility dimension to our knowledge base. Finally, PROTÉGÉ-II has a feature of generating knowledge acquisition tools from domain or application ontologies that allows knowledge experts specifies the domain concepts themselves. This aspect is as an add-on facility to our development process of knowledge bases, which gives flexibility to the experts and knowledge engineers to define new concepts and relations for the developed ontologies in this work.

Next, a history of ontology definitions used in various domains was presented, where in this work, we use ontology as a specification of a conceptualisation presented by Gruber [1995], which is the most referred definition in computing discipline. Our reasons for this choice are as follows. First, this view enables knowledge sharing and data reusing in coherent and consistent manner that is our ultimate goal of designing ontologies for the domains of study in this research. Second, this view of ontology associates the names of entities in the universe of discourse (for example, classes, relations, functions or other objects) with human-readable text. We take this advantage by defining a formal structure to describe the domain knowledge that is also understandable for our domain experts who do not have computational background. Finally, “conceptualisation” here is seen as an abstract and simplified view of the world where ontology engineers want to represent. Every knowledge base system or knowledge level agent is committed to some conceptualisation, explicitly or implicitly. This perspective is very critical for the domains of study in this research, that require explicit representation for various purposes at the application level.

Furthermore, among ontology representation languages discussed, we decided to use RDF and OWL languages to represent the ontologies in this thesis due to their capabilities of extendibility and computability through various interfaces. We also discussed three structures for ontology granularity, and presented a three-layer model of upper, domain and application ontologies. However, in designing the ontologies for this thesis, we did not use upper ontologies – SUMO and DOLCE – because these ontologies do not provide specific concepts for solving heterogeneity issues for the domains of study in this research.

In addition, we briefly discuss ontology learning. Since the primary step in any ontology
learning approach is the use of proper resources (e.g. existing domain ontologies, relevant Web documents and structured or semi-structured text), we have not considered this approach for the ontology development in this thesis because to our knowledge, there is no existing ontology or ontologies for the domains of sustainability and disaster management. Furthermore our text resources for these domains are mainly classified as unstructured text and we are unable to apply automated learning algorithms to extract the ontology components. Hence, our focus is to use ontology engineering methodologies.

We reviewed the importance of a reference ontology engineering methodology that provides a reliable guidance for developing ontologies. Several ontology engineering methodologies were discussed: Skeleton model, Formal method, On-To-Knowledge, SENUS-based approach and METHONTOLOGY, which were resembled to the Skeletal methodology [Uschold and King, 1995]. Each approach contains its own structure and some limitations which does not fully address our purpose of ontology engineering in this thesis. Accordingly as discussed, we use METHONTOLOGY, which is also based on skeletal methodology and satisfies our conditions in ontology development. The discussion was followed by introducing six types of ontology design patterns that are solutions to the modelling problems, which cannot be addressed in ontology engineering field. As one example, we use Value Partition ontology design pattern to address the design problems in constructing the ontology for sustainability indicator sets. Furthermore, we provided a summary of three well-known ontology evaluation techniques: OntoMetric, OntoClean and ROMEO. OntoClean is limited to manual-classification with meta properties. The downside of OntoMetric is that ontologies evaluated by this method cannot be extend to a similar domain of knowledge. These drawbacks lead us to use ROMEO for evaluating the ontologies in this research, because it overcomes such constraints by dividing the evaluation process into easy-approachable tasks.

Finally, the two last sections of this chapter provided a background study of taxonomies, domain and application ontologies that are used in the domains of sustainability and social media in the context of disaster management.
2.12 Summary

In this chapter, we discussed various approaches and methodologies that have been proposed in literature related to the research questions of this thesis on how using semantic technology can resolve the heterogeneity issues of data representation in the context of sustainability indicator sets and geotagging emergency tweets.

Having covered the background of ontologies and knowledge bases, in Chapter 3, we elaborate upon the research questions of the thesis, and briefly review the methods and techniques that are used to address them. Three contribution chapters then follow, which describe our solutions in details. The final chapter concludes the findings in this thesis and outlines the future directions.
Chapter 3

Research Design

*The scientist is not a person who gives the right answers, he’s one who asks the right questions.*

–Claude Lévi-Strauss

In Chapter 1, we present three research questions, that are directed to the overall objective of the broader project. To be more specific, this research showcases an effort to build a robust and reusable Ontology for Sustainability Indicator Sets (OSIS), by incorporating ontology engineering and ontology evaluation approaches into the integrated software of Circles of Sustainability (CoS) for design and evaluation purposes (RQ1 and RQ2). In addition, we propose how the use of existing and new ontologies can assist the disambiguation of geotagging tweets in the context of disaster management (RQ3). Furthermore, we develop a sustainability reporting software and a geotagger module to apply and evaluate our methods that can be used for different purposes of reasoning, applying and querying later.

This chapter describes the steps and methodologies used to address each question. The chapter begins with a discussion of a knowledge base design, for sustainability indicators in response to RQ1. It then describes the mechanisms used to browse heterogeneous indicator sets, in response to RQ2. Finally, it discusses the methods to resolve the heterogeneity of georeferencing in emergency tweets, to address RQ3.
3.1 A Knowledge Base for Representing Heterogeneity of Sustainability Indicator Sets

The first research question investigates the development and evaluation of ontologies that can be used to construct a knowledge base for representing semantic heterogeneity of sustainability indicator sets as articulated below.

**RQ1: How to develop an ontology for constructing a knowledge base that systematically represents semantic heterogeneity of sustainability indicator sets?**

We begin by constructing a knowledge base that represents semantic heterogeneity of sustainability indicator sets. As discussed earlier in this thesis, ontological approaches are chosen to construct the knowledge base by adapting an ontology engineering approach – a discussion of various ontology engineering methodologies was presented in Section 2.6, in this chapter we only describe the methodology that is used in designing the ontology of this research. Next, an ontology design pattern is used to tackle the modelling problems of ontology development. That is followed by constructing two ontology design candidates which are greatly different in terms of abstraction. The final step is applying an ontology evaluation technique that evaluates our two design candidates with the conclusion of the most appropriate ontology model as the input for the second phase of this research.

3.1.1 Adapting an Ontology Engineering Approach

To construct OSIS, we first developed an ontology engineering approach adapted from METHONTOLOGY [Gómez-Pérez et al., 2004], discussed previously in Section 2.6.5. Our adaptation – described more fully in Chapter 4 – emphasises or eliminates some of these activities using a set of rules:

- Depending on the management decisions and the purpose of the ontology, the association between support and development activities is diverse. For example, conceptualisation has a strong relationship with knowledge acquisition.
- Documentation has similar impact upon all development activities.
Some activities do presume dependencies; for example, there is typically more emphasis on the evaluation at the end of the implementation phase rather than on other activities.

3.1.2 Using Sustainability Indicator Sets

In developing OSIS, we use GRI and OECD indicator sets, which we now describe in greater detail.

The GRI reporting framework is designed to serve organisations’ economic, environmental, and social sustainability performance of any size, sector or location. The main advantage of the GRI reporting framework is the practical considerations faced by a diverse range of organisations from small enterprises to those with extensive and geographically dispersed operations. The GRI reporting framework guideline contains general and sector-specific content agreed by a wide range of stakeholders around the world.

The mission of the Organisation for the Economic Co-operation and Development (OECD) is to improve the economic and social well-being of people around the world. The OECD reporting on sustainability provides a forum in which government can work together to share experiences and seek solutions to common problems at the country level.

After applying an ontology engineering approach, we look into alternative mechanism(s) that can result in an abstraction design of OSIS for extendibility and reusability purposes. One of these mechanisms is the use of Ontology Design Patterns that we discussed in Section 2.7. Next, the design pattern used in this research is described.

3.1.3 Using an Ontology Design Pattern

We identify several modelling problems during the OSIS design – given in Section 4.2.1 – that lead us to use Value Partition (VP) pattern – described in Section ??.

The Value Partition ontology pattern is first introduced by Rector [2003] and is further reviewed and developed by Aranguren [2005] for the bio-medical domain. VP pattern represents specified collections of “values” – also known as “feature space” – using hierarchical modelling. Generally speaking, in any domain, such characteristics are used to describe different concepts in the ontology. Given the description of “Indicator Set” concept in the
sustainability domain. For example, given the description of “IndicatorSet” concept in the sustainability domain, in the presented ontology model, there are three values for the status of a indicator’s organisation: GRI, OECD and UN. To represent this feature space, two patterns are defined for the VP as follows.

- **Pattern 1** represents the `IndicatorSet` feature as a class, the feature space instead is a set of disjoint subclasses of the aforementioned organisations.

- **Pattern 2** represents the `IndicatorSet` feature as a class but the feature space instead is the individuals that enumerate the three values of organisations.

Our reasons for choosing these patterns are based on the advantages and disadvantages of two versions, that are compared by Rector [2003], given in Table 3.1.

### 3.1.4 Developing Two Ontology Design Candidates

Following the ontology engineering process of METHONTOLOGY and using the Value Partition pattern, we develop two ontology design candidates: Specific Ontology for Sustainability Indicator Sets, which we call it **SOSIS** and Generic Ontology for Sustainability Indicator Sets, which we call it **GOSIS**. The two design candidates differ largely in terms of abstraction as briefly describe below.

1. **SOSIS design**
   In this design, our emphasis is on the organisations that specifically develop sustainability indicators. We use pattern 1, that includes the key concepts of these organisations with their own indicator classifications. As a result, **SOSIS** design uses a range of classes and relations that are specifically added for each sustainability indicator set.

2. **GOSIS design**
   In this design, we use pattern 2, that defines broadly a suitable structure and reflects the generic key concepts of sustainability indicators. As a result, **GOSIS** design applies an object-oriented approach to encapsulate the generic features of all indicator sets.
Table 3.1: Advantages versus Disadvantages of Value Partition Pattern 1 and Pattern 2

<table>
<thead>
<tr>
<th>Value Partition</th>
<th>Pattern 1</th>
<th>Pattern 2</th>
</tr>
</thead>
</table>
| Advantages      | I Classes and subclasses can be subdivided into more partitions.  
                 | II There are also several alternative partitionings of the same quality space.  
                 | III Any subset of the OWL-DL and DL reasoner can classify the semantics of the partitions into set of discrete values correctly. | I OWL-DL and other used reasoners in DL is able to represent individuals correctly.  
                 | II From the database point of view, there is no need for the assumptions or conventions about anonymous individuals and the results can directly match to the correspondences in the databases. As a result, this approach seems more intuitive. |
| Disadvantages   | I From the database perspective, specific symbols are required to convert different anonymous instances of the classes to the meaningful schema in the database.  
                 | II Additionally, a list of features are often enumerated list of symbols in the databases, therefore, the use of classes as values is considered an “unintuitive” approach to many people from this background. | I Due to the OWL’s constraint of equality or differences between individuals, there is no possibility for individuals to have sub-partitioning values. In other words, an individual cannot be the union of other two individuals.  
                 | II Due to the previous limitation, it is not also possible to represent alternative partitioning of the same quality space. |

The development of design candidates is discussed in Section 4.2.1.

The final step of the OSIS development is evaluating the ontology. We discuss different ontology evaluation methods in Section 2.8. Next, the ontology evaluation method that is used in this research is described.

3.1.5 Applying an Ontology Evaluation Method

To evaluate the OSIS ontology, among ontologies evaluation approaches – discussed in Section 2.8 – we chose ROMEO [Yu et al., 2009].
For OSIS evaluation, we define an ontology role following by two requirements of precise and intuitive, which lead us to define two criteria questions of coverage and capture presented in Section 4.4. Given these components, we adapt precision and recall presented by Guarino [2004] to measure the coverage criterion. These metrics are originally given from Information Retrieval (IR) literature – known as retrieval performance evaluation [Baeza-Yates and Ribeiro-Neto, 1999] – that determine the quality of the answer set generated from a query task. In addition, they can be applied in the context of ontologies. For example, Rodriguez and Egenhofer [2003] use these metrics to measure the fraction of similar entity classes from different ontologies, and Euzenat [2007] also uses such metrics for measuring ontology alignments. We also apply the $F$-measure which is the harmonic mean of precision and recall that provides a sense of adequacy and balance modelling of the domain being presented in the ontology.

In Section 4.4, we describe how applying ROMEO and its tasks assist in evaluating the OSIS.

### 3.2 Mechanisms for Browsing Heterogeneous Sustainability Indicators

While the first question aims at expressing semantic heterogeneity of sustainability indicators at the data level through the use of ontological design, the second question explores various techniques of representing semantic heterogeneity of sustainability data for a sustainability reporting framework at the user interface level, that is phrased as:

**RQ2:** Assuming to have found a positive answer to RQ1, what mechanism can be used in a sustainability reporting framework to help end-users for navigating and browsing heterogeneous sustainability indicator sets?

To address this question, we first describe the development process of Circles of Sustainability (CoS) software that supports the associated model and methodology by applying the generic design of the OSIS given from the first research question. Second, we develop the user interface that provides two ways of browsing and navigating sustainability data: Tabular and
Circular views, which is followed by reviews of related work about user interface evaluation approaches and usability analysis measures. Finally, we conduct a user study to evaluate the perceived usefulness and perceived ease of use of the user interface that assists end-users to cope with the semantic heterogeneity of browsing sustainability indicators.

3.2.1 Structure of Circles of Sustainability Framework

In addressing the second research question, where we look at user interface approaches to dealing with semantic heterogeneity, we develop CoS software, which is adapted to the broader framework used by the United Nations Global Compact Cities Programme\(^1\). Circles of Sustainability framework is first introduced by Scerri and James [2010]. Its methodology is later presented by Magee et al. [2013], which consists of three phases including: (i) A domain model associated with conceptual entities (such as indicators and issues), (ii) A process for constructing the entities applying predefined rules (for example, an indicator must measure one or more issues), and (iii) A software for supporting the methodology. Each phase is described below.

**Circles of Sustainability Model**

The Circles of Sustainability model contains five main components:

1. Domains and subdomains.
2. Issue, normative goals and objectives.
3. Indicators, indicator sets and targets.
4. Network Relationships between issues, indicator sets and individual indicators.
5. Data collected against the indicators.

According to Magee et al. [2013], the first component consists of the four domains – Economics, Ecology, Politics and Culture – and associated subdomains illustrated in Figure 3.1. At the heart of these components is the concept of the issue, that is in between of overall

\(^1\)http://citiesprogramme.com/
reporting context (such as an organisation, or key organisation project), the domain model, and the specific variables or indicators against which data is collected. Various issues can be related to the entire reporting scope (general), or be identified with some level of attention (particular). A general issue can correspond to a shared intention agreed within the system (normative goal). A consensus of the stakeholders may also express some particular desirable state of affairs (objective) for a subordinate issue. The third component consists of a set of indicators and associated targets that are observable and measurable by issues. The fourth element includes establishing a set of relationships between the conceptual entities above. For example, one of the most fundamental relationship is suggested by Magee et al. [2013], that is the normalising an issue using some combinations of indicators. The final element
includes the data collection during the administration of the reporting framework.

Circles of Sustainability Methodology

The purpose of the Circles of Sustainability methodology is to specify several rules for constructing the entities of the model. The key steps to develop the methodology are as follows:

1. Definition of general issue and associated normative goal
2. Determination of critical issues and relevant objectives
3. Selection of indicators and associated targets
4. Establishment of relationships between domains and subdomains, issue sets, issue and indicators
5. Reporting against chosen indicators
6. Developing and monitoring a response
7. Reviewing and adapting an indicator model

Again following, Magee et al. [2013], the first step develops a general problem (a project) that is raised by one or more organisations and communities of stakeholders, which is followed by realising normative goals of the project. The second step identifies particular issues associated with the general problem, outputting a list of critical issues that helps in defining project objectives. The third step is built upon the previous results, which enables developing a set of indicators for measuring issues. Indicators may have been already defined by various organisations or can be created for a particular issue. There might be the need of supplementary relationships among components of the Circles of Sustainability model, such as domains, subdomains and indicators, which emerges the fourth step. At the end of this step, a sustainability reporting team should have four completed components: a general project definition and associated normative goals, a series of critical issues and relevant objectives, a database of indicators with associated targets and finally, a set of supplementary relationships between issues and indicators which are varied for different applications. An
example provided by Magee et al. [2013] is a relationship between the use of available product recycling approaches and additional cost pressures on waste disposal that causes explanatory power to the eventual reporting framework. During the fifth step, a sustainability report is produced through the specification of indicators types and rules. That is followed by steps six and seven at the end, which result in the final policy development, monitoring and adjusting the sustainability reporting framework to evolve issues. In this thesis, we use the first four steps to develop and evaluate the CoS software for browsing sustainability indicator sets. Steps 5 to 7 are not taken into the consideration because they are required for developing a specialised reporting framework in the domain of sustainability indicator sets, which is out of the scope of this research.

3.2.2 Developing Circles of Sustainability Software

In this research we design the Circles of Sustainability software as an associated Web-based software that embeds the aforementioned methodology and model to construct a meaningful indicator-reporting framework. It offers two ways of browsing semantic heterogeneity of sustainability indicators with the focus on methodology adoptable to multiple community and organisational reporting contexts. Its architecture contains three main layers of Database, Back-end and Front-end. Each layer includes various components that are described next.

Database

For the database, we use PostgreSQL 9.3.2\(^2\) that is an open-source object rational database system. It has full support of functions and data types used in SQL language. More importantly, PostgreSQL is compatible with RDF technology enabling the expression of structural forms. This is one of the main reasons for choosing PostgreSQL.

SPARQL-end

Jena SDB\(^3\) engine is also used as the RDF store and SPARQL endpoint. More details about RDF and SPARQL are given in Section 2.3.3.

\(^2\)http://www.postgresql.org/
\(^3\)http://jena.apache.org/documentation/sdb/
Back-end

The back-end of CoS software includes the components that support technical requirements of the system including: a lightweight framework for implementing the conceptual entities and an ontology model to define the relationships between entities. The components are as follows.

Play Framework

Play\(^4\) is an open source Web-application framework using a lightweight, stateless and Web-friendly architecture implemented in Scala\(^5\) and Java\(^6\). To compare with the standard Java APIs, it provides more predictable and minimal resource consumption (such as CPU, memory and threads) for highly-scalable applications. One of the reasons for this advantage is the use of Model-View-Controller (MVC) architectural pattern that aims to optimize the software development process by separating internal representations of information from ways of displaying them to users.

The MVC pattern, as shown in Figure 3.2, is used for implementing user interfaces by dividing the software application into three interconnected parts. The Model filters raw data including application data, business rules, logic and other functions, which causes programmer to think of filtered data as simple cognitive models such as objects. The View displays the Model in specific ways that favours the user request. This architecture also allows multiple views of displaying the same model in various forms. For example, one View may show data as a bar graph while another shows it as a pie chart. The Controller coordinates Views

---

\(^4\)http://www.playframework.com/
\(^5\)http://www.scala-lang.org/
\(^6\)http://www.oracle.com/technetwork/java/index.html
with Models. In other words, it plays as the communication channel to accept the input from keystrokes, locate device data and other events, and interpret them as commands for the Model to use.

The main reason for us to choose the Play framework is that it provides the representation of data with a more manageable way due to using MVC pattern that gives more flexibility to both programmer and user to compare with traditional way of the direct manipulation of data. Therefore, this framework satisfy our requirements for semantic heterogeneity of browsing indicator sets by suggesting more than one View.

Ontology Model

In what follows, we review how applying the generic ontology design (GOSIS) on the CoS software helps in defining relationships between the conceptual entities in the CoS model. The details of ontology engineering steps of developing GOSIS are described in Section 4.2. In designing GOSIS, our focus was presenting a generic model for sustainability data that covers broad key information of the domain and various indicator systems. The components identified in the ontology model correspond to the aforementioned conceptual entities of the CoS model – presented in Section 3.2.1. They are domains, subdomains, issues, indicator sets and indicators, in which “domains” and “subdomains” are replaced with Category and SubCategory due to various representations of the hierarchical structure of indicators in different sustainability reporting organisations. Description and Reference are also added for more clarification of the indicators.

As presented in Section 3.1.3, Pattern 2 of the Value Partition ontology design pattern constructs the basis of the generic design. In GOSIS, several sustainability reporting organisations are the values for IndicatorSet and an indicator belongs to one organisation at a time. From the point of view of reusability, indicator systems can be included as instances of IndicatorSet and Indicator class which are further linked together by a particular relation (such as belongsToIndicatorSet). In addition, two namespaces are used in this model:

- osis: refers to the OSIS URI\(^7\) used to represent the most abstract and generic concepts

\(^7\)http://www.cs.rmit.edu.au/sustainability/ontology
and relations, such as `<osis:hasIndicator>` and `<osis:hasDescription>`.

- `dc:` refers to the Dublin Core metadata URI\(^8\) to label common properties that pertain to most or all entities, for example `<dc:title>` for name entities and `<dc:type>` for type entities.

**Front-end**

The front-end of the CoS software is the interface of the Web page. We apply the Responsive Web Design (RWD) approach [Natda, 2013] that aims at providing an optimal displaying mechanism for easy reading and navigation experience through the Web pages with a minimum of resizing, panning, and scrolling across a wide range of devices from mobile phone screens to desktop monitors.

To put forward the CoS software interface, we also use Bootstrap 3\(^9\) which is a new and widely used customised technology of Web design that facilitates using CSS and JavaScript files by providing frequent suggestions of pre-made widgets and components. This technology and an associated light-weight theme were compatible with the Play framework from the back-end component.

As a whole, the components described above, form a unified architecture for the CoS software.

### 3.2.3 Browsing Heterogeneous Sustainability Indicator Sets

In pursuit of supporting the end-user to navigate heterogeneous sustainability indicator sets, two hypothetical mechanisms of browsing indicators are used. The displaying methods are the **Tabular** and the **Circular** views which are types of text-based and visual interfaces respectively. Text-base interfaces are easier for search entities because of more intuitive rendering and navigating features. Some examples are: Sig.ma [Campinas et al., 2011], SView [Cheng et al., 2013] and SWSE [Hogan et al., 2011]. Visual interfaces are used individually or in combinations of graphical structures such as images, maps, graphs and timelines to represent

---

\(^8\)http://purl.org/dc/elements/1.1/

\(^9\)http://getbootstrap.com/
information. Some examples are: VisiNav [Harth, 2010] and DBpedia Mobile [Becker and Bizer, 2009].

Our own evaluation compares the textual Tabular view and the visual Circular views. The implementation details of these are given in Section 5.1.2.

3.2.4 Evaluating User Interfaces

The purpose of usability inspection is to identify and address usability problems of an entire design. This topic has been increasingly applied as a method to evaluate user interfaces since 1990s. Evaluating user interfaces are classified into four categories:

1. Automatic evaluation by running the interface specification through some programs.
2. Empirical evaluation by testing the interface with real users.
3. Formal evaluation by applying exact models and formulaties.
4. Informal evaluation based on rules of thumb of evaluators’ experience.

While there is a little work on automatic and formal evaluation approaches due to their complexities and constrains for large scale interfaces, empirical approaches have received more attention in the literature with user testing probably the most commonly used method.

Two broad views for evaluating user interfaces are pointed out by Park and Hwan Lim [1999] including: assisting in design decisions and measuring the quality of use. The first perspective affects the designers’ decision making of possible alternatives (such as new user interfaces), and the second view reports on the value of some measures for comparing purposes, which results in emerging a field of comparative evaluation approaches for user interfaces. In comparative evaluation, some studies use multiple-criteria decision making to characterise the overall usability of an interactive system. Other studies apply an integrated assessment of several interface characteristics to measure the usability.

In addition, Davis [1989] studies the conceptual theories and empirical distinction of these variables from several divers lines of research, such as self-efficacy theory, cost-benefit paradigm, adoption of innovation, evaluation of information reports and channel disposition.
model. The user behaviour is also examined from the point of view of the vendor organisations such as IBM, Xerox, Digital Equipment Corporation, in which, user perception is influential in decisions to use in information technologies and it makes “usability testing” as a standard phase of product development cycle. One problem with widely used subjective measures in practice is that the quality of these measures are not validated to correlate with the user behaviour.

While evaluation approaches of user interfaces is a controversial topic itself, the definition of usability is also a pressing topic in the literature, that we discuss next.

**Usability Measures**

The International Organization int [1998] defines usability as the extent “to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction”. The common dimensions of usability, effectiveness, efficiency, and satisfaction, are used to elicit evaluative data from subjects about these qualities. A survey conducted by Hornbæk [2006] on current practice of Human Computer Interaction (HCI) research defines each dimension and lists several usability measures for each dimension. Effectiveness refers to accuracy and completeness of a tool which helps users to accomplish particular tasks. Some examples of effectiveness measures are: “Binary task completion”, measuring the number or the percentage of the tasks that user completes successfully, and “Quality of the outcome”, measuring the quality of users’ understanding and their learning of information provided in the interface. Efficiency measures the resources expended in relation to achieving the goals by the user. In other words, a software is efficient if it helps users complete their tasks with minimum waste, expense or effort. Most measures in efficiency are time-based as time is an important factor in assessing various tools. Satisfaction is defined as the positive attitudes of the user to the product. It also refers as the fulfilment of a specific desire or a goal [Kelly, 2009]. This metric aims at measuring gratification and contentment of the users experience when they accomplish particular goals. Some examples of satisfaction measures are: “Preference”, measuring satisfaction as the interface users prefer using and, “Ease of learning”, measuring the amount of effort expend learning to use a system.
We use satisfaction dimension to define the variables in the questionnaire of the user study – given in Section 5.2.2 – because satisfaction is a significant indicator for measuring the system complexity from the user perspective and user perspective towards the CoS software is the main focus of evaluating user study in this research. In other words, satisfaction is a broad measure of users’ overall satisfaction or attitudes towards the interface or user experience. Particularly, we focus on two specific measures of perceived usefulness and perceived ease of use, that are the fundamental concepts of user acceptance presented by Davis [1989]. Perceived usefulness is defined as the degree to which a person believes that using a particular system would enhance his or her job performance, which is extracted from the definition of the word useful that is referred as “capable of being used advantageously”. In a parallel argument, the author describes perceived ease of use as the degree to which a person believes that using a particular system would be free of effort, which also follows from the definition of ease meaning “freedom from difficulty or great effort” [Davis, 1989].

3.2.5 Conducting a User Study

A complete review of methods for evaluating information retrieval system with users is presented by Kelly [2009]. Kelly identifies the following six types of evaluation studies including: Descriptive Explanatory; Evaluations and Experiments; Laboratory and Naturalistic; Longitudinal; Case Studies; and Wizard of Oz Studies and Simulations.

Among these, we chose a user study (survey) approach to evaluate the user interfaces of the CoS software, given in Section 5.2. The aim of the evaluation was to gain feedback from users after they have used the software. Studying several respondents’ answers to questions and their statements helped us to generalise about the beliefs and opinions of other potential users and as a result their feedback assisted us to improve our method.

3.3 Resolving Heterogeneity of Georeferencing Emergency Tweets

The final research question addresses how the use of existing and new knowledge bases can resolve the semantic heterogeneity of georeferencing in tweets’ content in the context of disaster management, that is phrased as:
RQ3: How to use knowledge bases to automatically geotag emergency tweets which contain semantic heterogeneity in their location?

To address this question, we first introduce the OzCT geotagger which is developed for the OzCrisisTracker application. We then review the studies that process geographic references. It is followed by a summary of related work using Toponym Recognition and Toponym Resolution approaches, which are used in OzCT geotagger. Next, several geographic knowledge bases applied to enrich OzCT geotagger are discussed. Finally, we describe the evaluation process of our geotagger against two other methods.

### 3.3.1 Developing OzCrisisTracker Geotagger Module

OzCT geotagger module is one of the components of the OzCrisisTracker Web-application.10 This application aims to provide an overview of significant natural disaster events in Australia from Twitter messages. Given the scarcity of user geotagged tweets, the aim of the OzCT geotagger is to automatically identify and reference geolocations in the tweet contents so that different geographic events can be easily identified. This helps with clustering tweets that relate to same geographic event and thereby aids in reducing information overload. In addition, tweet geotagging can help in identifying events that span a larger geographical area and hence signal an unfolding disaster [Ikawa et al., 2013; Vieweg et al., 2010].

### 3.3.2 Processing Geographic References

A brief review of existing research is provided in this section. A complete survey on geotagging strategies can be reviewed in a PhD thesis by Leidner [2007].

*Geotagging* (also known as *geocoding*) refers to a process of identifying phrase portions in unstructured texts with possible spatial aspects and disambiguating these references by linking them to location information typically geographic coordinates. Social media messages contain different types of locations such as a place name, an incomplete address or a city

---

10http://www.ozcrisistracker.com
or suburb name. One of the main challenges in geotagging discussed by Lieberman et al. [2010], is the ambiguity involved in the names of the location used in unstructured text (called toponyms). Amitay et al. [2004] tackle this problem by resolving two types of ambiguities for Web pages: geo/non-geo and geo/geo, in which, geo/non-geo ambiguity is distinguishing non geographic entities in the text and geo/geo ambiguity is identifying distinct places with the same name. Similarly, Lieberman et al. [2010] present toponym recognition and toponym resolution for disambiguating location references in news articles, in which, during the first phase all toponyms are identified and in the second phase each toponym is assigned to the correct geographic coordinates among the possible candidates.

The latter technique is our initial motivation in this research.

**Toponym Recognition**

At the phase of toponym recognition, geo/non-geo entities are distinguished to decide whether an entity refers to a location or some other entity such as a person’s name. Lieberman et al. [2010] classify two general approaches for toponym recognition. The first and the most common strategy is finding similar phrases in the document that exist in the gazetteer, which is an external database of geographic locations. The phrase matching could be performed at various levels, depending on the size of the gazetteer. For instance, Web-a-Where [Amitay et al., 2004] uses a small, well-curated gazetteer of about 40000 locations created by collecting countries and cities with populations greater than 50000; whereas, Lieberman et al. [2010] use a gazetteer (GeoNames) with nearly 7 million entries to recognise the local toponyms. One of the disadvantages of using larger gazetteers is the ambiguity in recognising toponyms. Several heuristics for filtering potentially erroneous toponyms have been proposed [Rauch et al., 2003; Stokes et al., 2008].

The second type of approach in toponym recognition is rooted in solutions to related problems in Natural Language Processing (NLP), namely Named-Entity Recognition (NER) and Part of Speech (POS) tagging. The main disadvantage of such techniques is pointed out by Stokes et al. [2008], which is statistical NER methods can be useful for analysis of static corpora but they are not well-suited to the dynamic nature of Web resources such as news.
To avoid such complexity, hybrid approaches are proposed by Lieberman et al. [2010].

**Applying Tweet Toponym Recognition**

Our approach uses a technique similar to Lieberman et al. [2010]. This method first determines *local lexicons* within a collection of articles in an online news resource by applying a hybrid toponym recognition technique utilising GeoNames\textsuperscript{11} gazetteer. The ambiguity of the identified local lexicons are then resolved by a set of heuristic rules. These rules can be established easily by using linguistic contextual clues, for example, news audiences read articles linearly and as a result, article language has a contextual and geographical flow. However, our method is different from this approach from various aspects, that is described in Section 6.2.

**Toponym Resolution**

A toponym resolution procedure resolves the *geo/non-geo* ambiguity, once all toponyms are recognised. The simplest strategy is to assign a default sense to each recognised toponym using prominence measures such as population [Stokes et al., 2008; Rauch et al., 2003]. Another popular technique is “resolving context within a hierarchical geographic ontology”. The main assumption behind this approach is that the considered document has a single *geographic focus*. If there is more than one geographic focus, other techniques are proposed such as setting a node with the highest score in the hierarchical ontology [Amitay et al., 2004], selecting the most frequent toponyms [Ding and Gravano, 2000] and heuristic techniques discovered by humans [Lieberman et al., 2010] which are also used in our method.

**Applying an Ontological Model**

A further step in toponym resolution, we also develop an ontology model to apply semantic annotation of Twitter messages, that uses existing and new ontologies to improve further investigation of resolving the ambiguity of the location results and referencing geographic coordinates within the tweets’ content. The details of this model are given in Section 6.5.2.

\textsuperscript{11}http://geonames.org
3.3.3 Using Geographical Knowledge Bases

In the context of the georeferencing, a geographical knowledge base refers as a *Gazetteer*. A gazetteer is an external geolocation database which contains a set of geographical entities associated with geographical coordinates. Some researchers build their local gazetteers utilising existing data resources while others use a dynamic approach to query location entities from global geographical containers. With regards to the former technique, Amitay et al. [2004] develop a hierarchical ontology which associates each geographical entity with a canonical node. In another example Lee et al. [2013] build a dataset of *venues* and their associated properties (*geolocation tips*) by utilising FourSquare’s public API, which is a mobile application that allows individuals to *check-in* at a particular venue and leave tips for that venue. One instance of the latter technique is the system presented by Yin et al. [2012] that detects and passes the location string identified in the tweet content to the Yahoo geocoding service\(^\text{12}\) in real-time and retrieves the top five matches.

Our approach is a combination of both techniques. We first use the training dataset – given in Section 6.1 – to identify location entities and dynamically query them through public data sources including Google Map API\(^\text{13}\) and GeoNames Dataset\(^\text{14}\) while narrowing the search criteria to Australian location entities only. We then store the retrieved geographical references into a table in a relational database (IBM-DB2\(^\text{15}\)) to be used as a static map later.

3.3.4 Evaluating the OzCT Geotagger

We evaluate the performance of the OzCT geotagger using two techniques. First, we evaluate the system by comparing its output against human judgement as ground truth for a selection of 500 tweets (containing simple and complex geolocation references) that were manually geotagged. We choose manual geocoding as the first evaluation technique, because it is widely used in other research. For example, the Open Directory Project (ODP)\(^\text{16}\) is maintained by a vast number of volunteers who have created the largest human-edited directory of the

\(^{12}\)http://developer.yahoo.com/boss/geo/
\(^{13}\)https://developers.google.com/maps/
\(^{14}\)www.geonames.org/
\(^{15}\)www.ibm.com/software/data/db2/
\(^{16}\)http://www.dmoz.org/
Web (3.8 billion pages) and a random sample of 20,000 Web-pages from the ODP’s Regional section is used to evaluate the focus-finding algorithm in the Web-a-Where project [Amitay et al., 2004]. Similarly, Ikawa et al. [2013] create a gold standard dataset of all places that were manually extracted from the Twitter messages to evaluate the precision and recall.

Second, we test our geotagger performance with the testing dataset against other geocoding platforms: the Alchemy API Named Entity Recognition and the Yahoo! Geo-Planet service, described in Section 6.4.2. There are two reasons why these services are chosen specifically. (i) To best of our knowledge, in the current literature, there is almost no package for geotagging tweets that we could use to compare against our geotagger. (ii) These platforms are widely used in different research with various purposes, which make the tweet geotagging process non-biased. The details of the evaluation techniques are given in Section 6.4.

3.4 Summary

The aim of this thesis is to develop and evaluate ontological solutions to resolve semantic heterogeneity of sustainability indicator sets with the development of a software for the CoS sustainability reporting framework and disambiguating georeferences of tweets’ locations for the OzCrisisTracker application. In this chapter, we have described three research questions along with the related methodologies that are used to address each question. Chapter 4 addresses the first research question, developing two ontology design candidates by adapting an ontology engineering approach and applying an ontology evaluation technique. Chapter 5 describes our method to address the second research question, developing the CoS user-interface by applying the generic ontology model and conducting a user study to evaluate two methods of browsing and navigating heterogeneous data of sustainability indicators. Finally, Chapter 6 explores third research question by developing a geotagger for tweets in the context of disaster management, in which, the use of a semantic annotation technique resolves the disambiguation of georeferences.
Chapter 4

An Ontology for Sustainability Indicator Sets

A thinker sees his own actions as experiments and questions, as he attempts to find out something. Success and failure are for him answers above all.

–Friedrich Nietzsche

Sustainability indicators are used to measure the current and estimated future status of complex systems, such as cities, organisations and ecosystems. In this context, a “system” is referred to as a framework or an organisation in one of the domains of culture, economics, or ecology that performs various tasks including reporting, querying and assigning sustainability indicators. In response to measuring and maintaining the sustainability status of a system, several indicator systems have been developed and are in use today [Brilhante et al., 2006; Kumazawa et al., 2009; Pinheiro et al., 2009]. The diversity of indicator systems itself poses problems – for would-be reporting organisations – for how to find which representation of indicators best reflect the sustainability challenges. Sustainability challenges are therefore mirrored by semantic challenges due to the existence of multiple overlapping representations of the domain. These challenges remain at two levels: (i) Some complexities belong in naming the issues related to the indicators, which various organisations define in different ways, and (ii) Some challenges belong to the cognitive nature of the sustainability data. To
date, there have been few efforts to represent *multiple* indicator systems in a systematic way. In particular, we see considerable applications for the representation of indicators in a formal *ontology*. Through use of such a formal construct, it is possible to develop a consistent definition of what an indicator is, and how it can be applied. An ontology in turn would allow organisations to browse and review different kinds of indicators for different measurement applications, and to enable some degree of comparison and bench-marking between them.

A further challenge exists in the design and evaluation of domain-level ontologies. Although in ontology research several approaches have been proposed for structuring knowledge into different levels of abstraction [Chandrasekaran and Johnson, 1993; Richards, 2000; Steels, 1993] there is still considerable ambiguity involved in the processes of knowledge base construction and evaluation. As suggested in Chapter 3, promising approaches such as METHONTOMETRY [Gómez-Pérez et al., 2004] and ROMEO [Yu et al., 2009] are developed to provide clear structural guidelines, in order to simplify the complexities and eliminate doubts in ontology design and evaluation. However these methods are not yet deployed in the context of sustainability indicators, in the manner this research aims to address.

The current chapter showcases an effort to build a robust and reusable knowledge base for representing the semantic heterogeneity of sustainability indicator systems using an ontology, which is the first focus of this research. It is called OSIS, that stands for an *Ontology for Sustainability Indicator Sets*. We show the activities for constructing a knowledge base by adapting a well-known ontology engineering approach that is suitable for OSIS development. Ontology Engineering refers to a set of principles that are required for the ontology development process and the ontology life cycle. There are comprehensive ontology engineering methodologies – see Section 2.6 – which emphasise various aspects of ontology development depending on particular frameworks. Among various ontology engineering approaches, we choose METHONTOMETRY for OSIS development due to its high rate of adoption among other domain ontologies. A review of the METHONTOMETRY approach and our reasons for choosing this method is discussed in Section 3.1.1. Here, we have separated the activities of METHONTOMETRY into three phases: Pre-Design, Ontology Design and Post-Design. As shown in Figure 4.1, each phase includes two activities and the output of one activity is the
input of the next one. The present adopted ontology engineering method makes it possible to modify the previous phases at any stage.

The structure of this chapter is as follows. Section 4.1 outlines the activities of the Pre-Design process of OSIS construction containing specification and knowledge acquisition phases which result in determining the key information about sustainability indicator sets. Section 4.2 demonstrates Ontology Design activities, which develop ontology components, taxonomy and design patterns resulting in the construction of two ontology design candidates: generic and specific. We then outline the Post-Design processes of building knowledge base and step through ontology evaluation tasks in Sections 4.3 and 4.4 respectively. Section 4.5 describes the experiments in two stages. The results are presented in Section 4.6. Finally, we summarise and conclude this chapter in Section 4.8.
Purpose: The aim of the ontology is to represent knowledge about sustainability indicators in the context of a sustainability reporting application. Two specific requirements need to be supported by the ontology, which may not be necessarily addressed with one ontology design model.

1. It should be sufficiently *generic* and *reuseable* so that multiple indicator systems can be easily supported.

2. It should reflect the key concepts and structure of supported indicator systems *precisely* and *intuitively*.

<table>
<thead>
<tr>
<th>Level of Formality:</th>
<th>Formal - Represented in OWL.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scope:</td>
<td>Sustainability Indicator sets.</td>
</tr>
</tbody>
</table>

*Table 4.1: Ontology Requirements Specification Document for OSIS*

4.1 OSIS Pre-Design Process

While METHONTOLGY does not mandate a specific order, prior to beginning the formal design of OSIS we have undertaken *Specification* and *Knowledge Acquisition* activities as the key inputs to the Design phase.

4.1.1 Specification

The purpose of the specification activity is to produce an Ontology Requirements Specification Document (ORSD) in natural language describing informal, semi-formal or formal representation of an ontology. In other words, the ORSD identifies the scope, purpose and requirements of the ontology. It also specifies the level of formality required for the ontology, depending on whether terms and their meanings of the specific domain need to be codified in natural or formal language. Table 4.1 describes the ORSD of the OSIS ontology.

The two explicit requirements are derived from different domain experts and Circle of Sustainability project sponsors (see Section 3.1). These requirements are not necessarily mutually compatible and led us to identify two perspectives in developing OSIS. The first view, which has been term GOSIS, or “Generalised OSIS, supports the design decisions that are generic and reusable in various indicator systems – whether known or unknown at the time of the analysis – whereas the second view, which has been termed SOSIS or “Specific OSIS,
OSIS Pre-Design Process

captures the design decisions which are specific and intuitive for particular sustainability indicator sets. These points of view are taken into account in developing two ontology design candidates for OSIS in Section 4.2.1.

1. Relevance
This indicator seeks to identify the reporting organisation’s ability to use recycled input materials.

2. Compilation
2.1 Identify the total weight or volume of materials used as reported under EN1.
2.3 Report the percentage of recycled input materials used by applying the following formula:

\[ \text{EN2} = \frac{\text{Total recycled input materials used}}{\text{Input materials used}} \times 100 \]

3. Definitions
Recycled input materials
Materials that replace virgin materials that are purchased or obtained from internal or external sources,

4. Documentation
Potential information sources include billing and accounting systems,

5. References

Figure 4.2: A Sample GRI Indicator

4.1.2 Knowledge Acquisition

In order to properly translate the specification into design, we have conducted a knowledge acquisition activity which is pivotal to develop a suitable domain ontology according to Gómez-Pérez et al. [2004]. First, we consult our collaborative researchers from our broader project (see Section 3.1) through a series of interviews and workshops. Second, we analyse
Environmental and resource productivity
Carbon & Energy Productivity
CO2 Productivity

Notes & definitions


The very high per capita emissions of Luxembourg result, to a large degree, from the lower taxation of gasoline and diesel oil compared to neighbouring countries.

Sources
  www.iea.org/co2highlights/co2highlights.xls

Figure 4.3: A Sample OECD Indicator

several available sources of domain knowledge, including widely used indicator sets – the GRI and the OECD – in order to extract key domain concepts of the ontology. The details of these indicator sets are given in Section 2.9.1.

Reviewing both indicator sets as well as interviewing and discussing with the sustainability domain experts lead us to identify the key concepts of sustainability indicators highlighted in Figures 4.2 and 4.3 as follows.

- **Indicator**: An entity to represent the status of a social or natural system, e.g. EN2 (Figure 4.2).

- **IndicatorSet**: A collections of indicators that are typically developed by an organisa-
tion or community, e.g. GRI.

- **Category**: A group of indicators from an indicator set that are classified for a particular area, e.g. *Performance Indicators* from the GRI indicator set.

- **SubCategory**: A small set of indicators that are defined to measure a particular group, theme or aspect, e.g. *Materials* from the above category.

- **Issue**: The real-world problem or issue that is at risk and needs to be monitored by one or more indicators. e.g. *recycled input materials* from *EN2*.

- **Description**: The description about an indicator that is provided by the organisation, e.g. *Relevance, Compilation, Definitions and Documentation*.

- **Reference**: The references about the indicator and its description that are presented by various organisations, e.g. *Sources and Information*.

The proposed key terms are especially helpful in representing the naming issue of semantic heterogeneity of the indicator sets. These key informations cover a broad range of indicator sets that have various representations of sustainability indicators and address one of the aspects of the first research question.

### 4.2 OSIS Design Process

Having specified requirements and a developed set of key terms, we begin to design OSIS with the focus on two main phases suggested by the METHONTOLOGY approach [Gómez-Pérez et al., 2004]: *Conceptualisation* and *Formalisation*. These activities have moved us from a generic and abstract model of the domain through to progressively more specific design. During Conceptualisation, we have decided upon the need to develop two distinct conceptual models, each reflecting one of the two main requirements identified during the Specification phase. These two models are then formalised and implemented in the next activities of the OSIS development process.
4.2.1 Conceptualisation

Conceptualisation is the most important activity in ontology development. The outcome of this activity is a specification of the ontology components which should reflect a set of terms produced in the Knowledge Acquisition phase. In what follows, the OSIS ontology components and taxonomy are discussed.

OSIS Ontology Components

Using the key concepts identified in the Knowledge Acquisition activity, the OSIS ontology components are defined below.

The concepts are distinguished between those which are broad and abstract, for example, IndicatorSet and Indicator which we model as superclasses in one design candidate of OSIS, and those which are specific and thus form the subclasses in another design candidate of OSIS. Superclasses are inherited by the subclasses using is-a relation. Additionally, classes may have attributes that complement them. Examples of attributes are PeriodOfTime and Publisher. Relations express the interaction between two classes of associates with has-a relation, such as hasCategory and hasReference. Individuals include all indicators of the sustainability systems, such as EC1 and EC2 in GRI and 1.1. Production Based CO2 Productivity and 1.2. Demand-based CO2 Productivity in OECD. In addition, indicator sets can be represented as constants based on which ontology design candidate – specific or generic – is chosen. We skip the step of defining axioms due to the small size of our ontology. The next step after identifying the ontology components is developing the taxonomy.

OSIS Taxonomy

We continue by developing OSIS taxonomy, which represents a hierarchical structure of is-a relations between concepts. In a taxonomy, there is no distinction between the ontological status of conceptual entities, for instance, whether they ought to be represented by classes or individuals, or related by generic attributes or specific subclass relations. Given the GRI and OECD taxonomies presented in Section 2.9.1 and related Figures 2.4 and 2.5, we notice
that the key concepts of sustainability development in GRI indicator set are presented in six standard categories (*disclosures*) and their various subcategories (*aspects*). Indicators are at the bottom of this hierarchy. Similarly, the OECD indicator set illustrates the key information in five main categories that include two levels of subcategories (*themes*). Like GRI, indicators are at the bottom of the hierarchy structure.

The **OSIS** taxonomy is based on the hierarchical structure of the GRI and OECD indicator sets, in which *themes* and *aspects* are replaced with generic concepts of **category** and **subcategory**. The indicators are also represented at the bottom of the hierarchy.

On the other hand, in designing **OSIS**, we emphasise interoperability and reusability features which give the higher level distinction of concepts which help one to understand the concepts at the lower level. Additionally, since ontologies and knowledge bases are not widely used in the domain of sustainability, these requirements account for the most important features for any developed ontology associated with this area. In developing **OSIS**, we address these with the use of ontology design patterns (ODPs). A complete literature of ODPs is presented in Section 2.7. The ontology patterns that are used in **OSIS** structure are presented below.

**Ontology Design Patterns Used in OSIS**

Given the identified key concepts and relations and **OSIS** taxonomy, the relation between abstract concepts may have various interpretations from different perspectives. As illustrated in Figures 4.2 and 4.3, we notice that specific indicator systems, taken from organisations such as GRI and OECD, should be specified in relation to abstract concepts of **IndicatorSet** and **Indicator**. In other words, one of the design problems is the association of **Indicator** with **IndicatorSet**. This affects the relations of other concepts such as **Category**, **Description** and **Reference**. The question here is how to determine the relations between **IndicatorSet** and **Indicator** concepts to be represented as classes, subclasses or instances? Solving such design problems ideally should reflect the requirements of the final ontology design.

To address the aforementioned modelling problems, we use the **Value Partition** (VP) pat-
patterns, described in Section 3.1.3. Two possible ontology models following the VP patterns are detected.

- From the viewpoint of explicitness, it is usefulness to specify particular system indicators as subclasses GRI of a superclass (for example IndicatorsSet) (Pattern 1). This design supports a specific representation of the domain problem and reflects detailed views of each system indicator. This view is more detailed to include direct references to specific indicator sets, which is called Specific Ontology for Sustainability Indicator Sets or SOSIS. Figure 4.4 features our solution for the second design problem.

- From the viewpoint of reusability, we also see that system indicators can be included as instances of IndicatorsSet and an Indicator class is further linked by a particular relation (for example belongsToIndicatorSet) (Pattern 2). Figure 4.6 illustrates our solution for the second design problem. This view is more broader to cover sustainability indicators’ key information with no reference to any particular organisations and is called Generic Ontology for Sustainability Indicator Set or GOSIS.

Another way to compare the two design candidates is to look at their UML diagrams
Figure 4.5: UML Diagram of SOSIS Design Using Value Partition Pattern 1
which are built upon the aforementioned Patterns of VP shown in Figure 4.5 and 4.7 for SOSIS and GOSIS designs respectively.

Furthermore, the rationale of the two design candidates are in relation to the requirements defined in the Specification activity (Section 4.1.1). We are unable to present a unique ontology model, which addresses both requirements. Instead, we suggest the use of two ontology design models, in which each model captures an aspect of the requirements: Supporting indicator systems (i) precisely and intuitively (SOSIS) and (ii) generally and reusability (GOSIS). The details of formalising the design candidates are as follows.

4.2.2 Formalisation

This activity involves the transformation of the conceptual models defined previously into a formal representation of an ontology. Namespaces are used during the formalisation activity to distinguish similar attributes and relations used in various ontologies from each other. This involves explicit representation of concepts and relations as classes, attributes and individuals with the development of namespaces for grouping related entities. Our intent in this phase is to use the formal notion of entities defined in existing namespaces or create new namespaces to represent the list of concepts, relations and attributes identified for two ontology design
Figure 4.7: UML Diagram of GOSIS Design Using Value Partition Pattern 2
candidates. The namespace declaration in OSIS is represented in XML language, where an alias is associated with the URI of a conceptual resource:

- **osis**: refers to the URI\(^1\) used to represent the most abstract and generic concepts and relations, such as `<osis:hasIndicator>` and `<osis:hasCategory>`.

- **dc**: refers to the Dublin Core metadata URI\(^2\) to label common properties that relate to most or all entities, for example `<dc:title>` for name entities and `<dc:type>` for type entities.

- **gri**: refers to the Global Reporting Initiative URI\(^3\) for the properties that are specifically related to GRI sustainability indicator sets, such as `<gri:hasAspect>` and `<gri:hasDocumentation>`.

- **oecd**: refers to the Organization for Economic Co-operation and Development URI\(^4\) for the properties that are specifically related to OECD sustainability indicator sets, for instance `<oecd:hasDefinition>` and `<oecd:hasTheme>`.

The latter two namespaces are not used in the GOSIS design, whereas all namespaces are applied to the SOSIS design. In other words, where there is a need for a generic relation with no emphasis on particular indicator sets, such as `hasCategory` and `hasIndicator`, `osis` namespace is used. Whereas, the necessity for defining specific relations for a particular sustainability system in GOSIS requires us to use unique namespaces, such as the relations `<gri:hasAspect>` and `<oecd:hasDefinition>`.

### 4.3 OSIS Post-Design Process: Building the Knowledge Base

The first step of the post-design process in our ontology engineering model – given in Figure 4.1 – is building the knowledge base which represents the two design candidates of OSIS described earlier in this chapter. The output of this phase is the OSIS knowledge base, which is evaluated against seen and unseen indicator sets. The steps are outlined below.

\(^1\)http://www.cs.rmit.edu.au/knowledgebase/sustainabilityIndicatorSets/ontology/
\(^2\)http://purl.org/dc/elements/1.1/
\(^3\)http://www.globalreporting.org/
\(^4\)http://www.oecd.org/
4.3.1 OSIS Language Representation

Among various ontology languages reviewed in Section 2.3.3, we use OWL-DL in formulating two design candidates of OSIS. Our reason is that its full expressive capability and type separation capability captures complex semantics of concepts and relations in the sustainability domain in finite time. While no axioms are included in the two candidates at this stage, the choice of OWL-DL means that they can be easily supported in future. This is important for expressing equivalence between two individual indicators, for example. We generate RDF triples of our ontology design models. An example is shown in Figure 4.8. This process involves converting each concept and associated relations and attributes into equivalent semantic triple statements. Each subject (for example IndicatorSet class or an individuals) is linked to an object (for instance Indicator class or an individual) by a predicate (for example hasIndicator relation). Furthermore, we also use SPARQL as the query language, once the ontology is constructed and populated with sustainability indicator data. SPARQL provides suitable constructs for retrieving RDF data, and displays results as RDF graph.

4.3.2 Ontology Data Storage

For the triple store framework of RDF, we choose PostgreSQL\(^5\) due to its compatibility to both relational (SQL) and semantic (RDF and OWL) languages. More technical details of PostgreSQL was given in Section 3.2.2.

At the final stage, we convert the two ontology design candidates into RDF/OWL form using Protégé ontology editor\(^6\). Subjects are considered as the Domain concepts of the attribute or relation and objects are considered as the Range. We then populate GOSIS and SOSIS with the GRI and OECD sustainability indicators which form the knowledge base of

\(^5\)http://www.postgresql.org/
\(^6\)http://protege.stanford.edu/
4.3.3 Structuring OSIS with the Empire Model

We map the ontology model to the RDF model with the use of the Empire model, that is an implementation of the Java Persistence API \(^7\) for representing Semantic Web technologies like RDF and SPARQL. In addition, the Empire model provides an easy transition for accessing RDBMS through JDBC drive.

The Empire model controls the mapping between a Java bean and a RDBMS using the common annotations of JPA. Given the declared annotation of the JPA entity below,

```java
@Entity
public class Indicator
```

The Empire model also extends this method by adding an additional annotation to the class to specify its type:

```java
@Namespaces({"osis", "http://www.cs.rmit.edu.au/knowledgebase/osis#"})
@RdfClass("osis:Indicator")
public class Indicator implements SupportsRdfId
```

Where, the additional optional annotation, @Namespaces facilitates the use of specific keywords instead of full URIs. It is also essential to assert that the Indicator class have a RDF identifier, which is auto-generated.

In Empire model, properties of each Java bean are specifically mapped to the corresponding property of an instance.

```java
@RdfProperty("dc:title")
private String title;

@RdfProperty("dc:periodofTime")
private Date periodofTime;
```

\(^7\)http://www.oracle.com/technetwork/java/javaee/tech/persistence-jsp-140049.html
With these simple additional annotations, the Java bean can now be used with the Empire model. As an example, the triple structure of the ontology concepts and their relations in the GOSIS model are shown in Table 4.2, where each relation links one concept (Domain) to another concept (Range) or a data type (such as String and Date).

The knowledge base is then used as the input for the final phase of OSIS development, the ontology evaluation, as follows.
4.4 OSIS Post-Design Process: Ontology Evaluation

As described in Section 3.1.5, among several ontology evaluation approaches, we choose ROMEO to evaluate OSIS. Our main motivation for choosing this method is that ROMEO focuses on the ontology evaluation requirements as various tasks. In other words, it breaks down a complex process into approachable and easy-implemented steps. In what follows, we step through ROMEO methodology containing five steps. The process starts by defining several frames of reference.

<table>
<thead>
<tr>
<th>—</th>
<th>1st Level</th>
<th>2nd Level</th>
<th>3rd Level</th>
<th>Indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRI:</td>
<td>6</td>
<td>33</td>
<td>-</td>
<td>78</td>
</tr>
<tr>
<td>OECD:</td>
<td>5</td>
<td>16</td>
<td>28</td>
<td>47</td>
</tr>
</tbody>
</table>

Table 4.3: Number of Sub-category and Indicators per Category for GRI and OECD Taxonomies

4.4.1 Frames of Reference

A frame of reference describes documents in a requirements specification, controlled vocabulary or structured dataset. Yu et al. [2009] define a frame of reference $F$ as, $F = < F_c, F_i, F_r >$, where $F_c$ is a set of concepts, $F_i$ is a set of instances and $F_r$ is a set of relationships, which is the union between the set of relations between concepts $F_{cr}$ and the set of relations between instances $F_{ir}$. Here we interpret “frame of reference” to be the sorts of knowledge sources solicited during the Knowledge Acquisition activity – the indicator systems themselves.

Since OSIS is designed to support two key requirements of intuitiveness and reusability, we choose three indicator sets to evaluate the ontology design candidates against each other. The first two – the GRI and the OECD given in Section 3.1.2 – are the sources used to construct the initial design, which are referred as the “seen” frames of reference. The third one – taken from the UN – is used to evaluate the ontology candidates of OSIS, which is described as the “unseen” frame of reference since neither candidate has any explicit entities drawn from it.

Each system reflects subtle yet distinctive features of how sustainability is conceptualised by their respective organisations. Hence using 2+1 frames of reference allows us to triangulate
4.4.2 Establishing the Ontology Role

In defining the role of OSIS, we investigate how this ontology will be used in a sustainability reporting framework in future. One of fundamental aspects in sustainability reporting practices is to suggest relevant indicators for a particular issue. We define the role of OSIS to enhance the suggestion process of indicators by solving representation issues of semantic heterogeneity of indicator sets discussed previously. The role of OSIS is shown in Table 4.4.

<table>
<thead>
<tr>
<th>Application:</th>
<th>An integration framework for systematic reporting on sustainability indicators.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ontology:</td>
<td>A formal ontology which contains several concepts and richly defined relations in the context of sustainability.</td>
</tr>
<tr>
<td>Role:</td>
<td>Enhance effectiveness of query expansion module in suggesting indicators for query tasks by solving representation issues of semantic heterogeneity of indicator sets</td>
</tr>
</tbody>
</table>

*Table 4.4: OSIS Role*

4.4.3 Ontology Requirements

Ontology requirements reflect a specific competency or quality of the ontology that can be obtained from existing ontology requirements or application requirements. For that, we use the same requirements defined during the Specification activity given in Table 4.1, which are also associated with the ontology role and purpose described in the previous section. As shown in Table 4.5, the first requirement is that the ontology candidate provides a *precise* and *intuitive* representation for the indicator systems. The second requirement is that its *reusability* allows other indicator systems to apply it easily and extendedly. Other conditions required to be addressed in this phase are also described in Table 4.5.

4.4.4 Criteria Questions

We specify two questions for the OSIS ontology candidates and ensure each question is answered with respect to both the seen (GRI and OECD) and unseen (UN) frames of reference.
### Requirement OR1:
Does the ontology provide a precise and intuitive representation of the indicator systems it represents?

### Requirement OR2:
Does the ontology allow for other indicator systems to be easily incorporated using existing concepts, properties and relations?

---

**Analyze:**

**OSIS**

**For the purpose of:**

Querying an ontology

**With respect to:**

Having rich definition of concepts and correctness of relations between concepts

**From the viewpoint of:**

Expert Panel of Social Scientists in Sustainability

**In the context of:**

Sustainability Indicators

**Motivation:**

A suitable ontology for sustainability indicators sets (OSIS) needs to make correct suggestions of indicators to the user queries. To address this goal, adequate level of coverage and correctness of the components of the ontology (concepts, relations and instances) are essential.

---

<table>
<thead>
<tr>
<th>Table 4.5: OSIS Requirements 1 and 2</th>
</tr>
</thead>
</table>

The criteria questions are listed below.

1. Do the ontology components (concepts, instances and relations) adequately cover the terms of the given domain?

2. Do the ontology components capture the terms of the given domain correctly?

The first question examines the *coverage* criteria and the second question determines the *correctness* feature of the ontology. In ROMEO [Yu et al., 2009] correctness is measured whether or not the right concepts, relations and instances have been modelled correctly according to the frame of reference. Similarly, coverage is assessed by whether or not the ontology sufficiently captures key concepts in the given domain.

4.4.5 Measures

In the final stage, Yu et al. [2009] suggested adopting a set of measures that are compatible with the ontology requirements which allowed us to answer the criteria questions. We adopt *precision* to measure to the correctness criterion, by determining the percentage of overlapping terms in an ontology $O$ that overlaps with the set of terms from a frame of reference $F$ (Equation 4.1). Additionally, *recall* is used to measure the coverage criterion, referring to
the percentage of overlap between a set of terms from the ontology and the frame of reference (Equation 4.2).

\[
\text{precision}(O, F) = \frac{|F \cap O|}{|O|} \tag{4.1}
\]

\[
\text{recall}(O, F) = \frac{|F \cap O|}{|F|} \tag{4.2}
\]

We also apply the $F$-measure to ontology evaluation in the context of indicating appropriate coverage of concepts in the relevant frame of reference (Equation 4.3).

\[
F\text{-measure}(O, F) = \frac{2}{\frac{1}{\text{recall}(O, F)} + \frac{1}{\text{precision}(O, F)}} \tag{4.3}
\]

A discussion of the origin of above measures is provided in Section 3.1.5.

4.5 Experiments

In evaluating OSIS, we conduct six sets of experiments: comparing the ontology documents generated from GOSIS and SOSIS design candidates with three frames of reference. The indicators are selected from the category of Economy from the three frames of reference (the GRI, the OECD and the UN indicator sets). The experimental design consists has two phases:

1. Preparing the indicator sets’ metadata and populating the ontology design candidates.

2. Obtaining the overlapping terms with the original documents.

4.5.1 Preparing and Analysing Sustainability Indicator Sets Metadata

The GRI organisation uses XBRL language to define its sustainability metadata in a taxonomy. This structured hierarchy captures the individuals reporting concepts as well as the relations between concepts and other semantic meanings in the original document.
Experiments

1. Percentage of materials used that are recycled input materials. (Core)

2. This Indicator seeks to identify the reporting organization’s ability to use recycled input materials. Using these materials helps to reduce the demand for virgin material and contribute to the conservation of the global resource base. For internal managers and others interested in the financial condition of the reporting organization, substituting recycled materials can contribute to lowering overall costs of operation.

3. 2.1 Identify the total weight or volume of materials used as reported under EN1. 2.2 Identify the total weight or volume of recycled input materials. If estimation is required, state the estimation methods. 2.3 Report the percentage of recycled input materials used by applying the following formula: EN2 = (total recycled input materials used / input materials used) x 100%

4. Potential information sources include billing and accounting systems, the procurement or supply management department, and internal production and waste disposal records.

5. Recycled input materials. Materials that replace virgin materials that are purchased or obtained from internal or external sources, and that are not by-products and non-product outputs (NPO) produced by the reporting organization.


Table 4.6: An Example of GRI Indicator (EN2) in XBRL format
Experiments

Table 4.7: An Example of GRI Indicator (EN2) in NTriples format
To prepare and analyse the GRI XBRL, we phrase the required data from this document using SAX functions including: start-document, start-element, end-element, character and end-document. These functions are applied to the label tags with conditioning the relevant attributes such as id, xlink:type and xlink:label. An example of a GRI indicator from the Environment category and its XBRL representation are shown in Table 4.6.

On the other hand, the OECD organisation does not present its indicators in a structured or semi-structured format. The hierarchy of indicators and associated information is represented in a pdf document. In order to use the OECD indicators, we first organise the dataset into an Excel document. Next, we convert the Excel document into a XML structure, defining an XML schema for the OECD dataset using the hierarchy structure and extracted key terms from the original dataset such as group, theme, topic, measurability and description. An example of OCED indicator and its XML code after the conversion are shown in Figure 4.3 and Table 4.8 respectively.

During this phase, the two design ontologies are populated through the use of the Protégé ontology editor by the structured information extracted from the GRI and OECD datasets. Finally, we export the RDF format from the final ontology documents in Protégé editor. An example of the generated file in NTriples is shown in Table 4.7. It shows the triple statements consisting of subject, predicate and object representing the same key information of EN2 shown in Table 4.6 using relevant namespaces, concepts and appropriate relationships in the GOSIS design. The relations in the ontology are shown in blue, such as description, format and isReferencedBy.

4.5.2 Obtaining the Overlapping Terms

The final step in performing the experiments is obtaining the overlapping terms between the original frames of reference and the ontology documents. To produce consistent results, a pre-processing stopping algorithm and Porter’s Stemmer technique [Porter, 1997] were applied to the original documents. We then used the algorithm presented by Broder [2000],

8http://www.w3.org/2001/sw/RDFCore/ntriples/
Table 4.8: An Example of OECD Indicator in XML format

<OECD>
  <Group>
    Environmental and resource Productivity
    <Theme>
      Carbon and Energy Productivity
      <Topic>
        CO2 Productivity
        <Issue>The estimates are affected by the quality of the underlying energy data.</Issue>
        <Indicator>
          1.1. Production-based CO2 productivity
          <Type>M</Type>
          <Notes>
            <Measurability>
              The very high per capita emissions of Luxembourg result, to a large degree, from the lower taxation of gasoline and diesel oil compared to neighbouring countries.
            </Measurability>
            <Description>
              Production-based CO2 emissions from fuel combustion Emissions calculated using IEA energy databases and the Revised 1996 IPCC Guidelines for National Greenhouse Gas Inventories. For example, some countries, both OECD and non-OECD, have trouble reporting information on bunker fuels and incorrectly define bunkers as fuel used abroad by their own ships and planes....
            </Description>
            <Reference>
              <Source>
              </Source>
              <Information>
                ECMT(2007), CuttingTransportCO2Emissions: ECMT, Paris ...
              </Information>
            </Reference>
          </Notes>
        </Indicator>
      </Topic>
    </Theme>
  </Group>
</OECD>
which determines the syntactic similarity between two documents. The first document was
the textual representation of the ontology, and the second was the frames of reference. Each
document is considered as a sequence of tokens, divided into contiguous subsequences called
shingles, of length \( n \), that is also known as \( n \)-gram. Broder [2000]’s algorithm compared the
sets of \( n \)-grams from two documents and calculates their resemblance value.

4.6 Results

In this section, we present and discuss the results from the experiments described in the
previous section. As shown in Table 4.9, we used the average \( F \)-measure with associated
metrics, the average recall and precision, to compare the two OSIS design candidates against
seen (GRI and OECD used) and unseen (UN not used) frames of reference.

The number of terms (\( |F| \)) between the three frames is different for several reasons. For
example, in a comparison with the GRI and the UN, while both frames distinguish between
culture, economics, ecology indicators, the GRI is directed largely towards corporate sustain-
ability reporting while the UN indicator set is aimed at measuring nation-level sustainability
development. The GRI therefore includes more economic indicators, while the UN empha-
sises social indicators. In addition, the UN set includes a fourth category of “institutional”
indicators, which inflates the overall indicator count. This consequently affects the number
of terms in each ontology \( |O| \) and overlapping terms \( |F| \cap |O| \) for each set of experiments.

The graph in Figure 4.9 features the \( F \)-measure for the results of coverage. Comparing
the numbers for the GRI and the OECD with the UN frames reveals similar coverage (approx.
65%) for GOSIS design. By contrast, the \( F \)-measure shows significant difference between the

| Frame of Reference | \(|F|\) | OSIS Ontology | \(|O|\) | \(|F| \cap |O|\) | Precision Ave | Recall Ave | F Measure |
|--------------------|-------|---------------|-------|----------------|-------------|-------------|-----------|
| GRI–Frame          | 2560  | GOSIS design  | 2280  | 1602           | 0.702       | 0.625       | 0.661     |
|                    | 2560  | SOSIS design  | 2309  | 2090           | 0.905       | 0.816       | 0.854     |
| OECD–Frame         | 986   | GOSIS design  | 802   | 590            | 0.735       | 0.598       | 0.659     |
|                    | 986   | SOSIS design  | 890   | 765            | 0.859       | 0.877       | 0.867     |
| UN–Frame           | 500   | GOSIS design  | 445   | 325            | 0.660       | 0.733       | 0.682     |
|                    | 500   | SOSIS design  | 303   | 247            | 0.494       | 0.315       | 0.307     |

Table 4.9: Precision, Recall and \( F \)-measure of OSIS design Candidates against Seen and Unseen Frames of Reference
two frames for SOSIS design: the GRI-Frame and OECD-Frame have large proportions of coverage (on average 85%) whereas the UN-Frame’s number declines significantly (30%).

![Figure 4.9: F-measure of Two Ontology Design Candidates, GRI and OECD are used in constructing the ontologies and UN is not used.]

### 4.7 Discussion

The findings from the previous section indicate the novelty of our ontology designs. The two candidates, generic GOSIS and specific design SOSIS, differ largely in terms of abstraction. The GOSIS design applies an abstract ontology design pattern. Here, for example, the concept IndicatorSet is defined as a class, while specific instances of indicators are treated as individuals which instantiate properties and relations of the IndicatorSet class. By contrast, the SOSIS design treats each indicator instance as a class as well. Accordingly, they inherit rather than instantiate properties and relations of the IndicatorSet class. This produces a much larger ontology that maps directly to the specific frames of reference that it is derived from. Accordingly, SOSIS design scores higher F-measure results against the seen frames of reference (the GRI and the OECD have been used in constructing this model). However, as our results show, GOSIS design produces a better F-measure score against an unseen frame of reference, such as the UN that has not been used in informing this model.

We conclude that the specific design is preferable where the domain requirements require a high degree of fidelity to seen frames of reference, while the generic design offers greater
reuse in contexts where unseen sets of indicators need to be added to the ontology in an *ad hoc* fashion. As sustainability indicators themselves continue to evolve, for this particular domain we argue the generic design is preferable.

### 4.8 Summary

In this chapter, we have introduced the field of sustainability indicator systems and argued that ontologies are well-suited for representing such systems formally. We adopted METHONTOLOGY, a well-known ontology engineering methodology, to guide the development of two pilot ontology candidates for this domain. We then applied ROMEO to evaluate the candidates. The evaluation consists of precision for measuring the correctness and recall to test the degree of coverage of indicators against two frames of reference. The first two frames of reference (GRI and OECD) were used to construct the specific design; the third (UN) was only used in the evaluation.

We argue that the GOSIS design presents a generic model for the ontology with no direct reference to any sustainability indicators. Actual indicators are assumed to instantiate, rather than inherit, from the *IndicatorSet* class. We view this accordingly as a more concise and generic conceptualisation of the domain. In contrast, SOSIS design presents a specific model, with more class references to particular sustainability indicators. This results in a higher F-measure against the seen frames of reference (GRI and OECD) – but because the specific wording of concepts maps directly to that frame, it performs more poorly against the unseen frame of reference (UN).

As a further investigation, we apply GOSIS design on a sustainability reporting framework that is developed in the next chapter. Chapter 5 addresses the second research question: how applying an ontology model (GOSIS for example) associated with the use of various ways of representing indicator sets can assist end-users to tackle the issue of browsing heterogeneous sustainability indicators.
Chapter 5

Evaluating an Interface for Sustainability Indicator Sets

No, our science is no illusion. But it would be an illusion to suppose that what science cannot give us we can get elsewhere.

—Sigmund Freud

The issues of semantic heterogeneity of sustainability indicator sets appear at two levels: data and interface. At the data level, the issues concern representing multiple sustainability indicator sets, in which various organisations use unstructured or semi-structured data using different vocabularies to represent common concepts and relations in the other domains of sustainability. The second type of heterogeneity problems appears in browsing and navigating sustainability indicators at the user interface level. The first research question of this thesis explored one of the computational solutions for resolving the heterogeneity issues at the data level; namely, to construct a unified model with the use of ontologies that captures specific or generic features of indicator sets. We suggest such an ontology model can then applied on a sustainability indicator sets reporting framework to provide end-users with a facility for representing indicator sets. This is followed by the second focus of this research – tackling the heterogeneity issues at the user-interface level and investigating which mechanisms can be used in a reporting framework to help end-users browse and choose appropriate indicators.
CHAPTER 5. EVALUATING AN INTERFACE FOR SUSTAINABILITY INDICATOR SETS

Figure 5.1: Login Page of CoS Software

from a range of heterogeneous sustainability indicator sets. In Chapter 4, we demonstrated the construction and evaluation of two design candidates – GOSIS and SOSIS – for the ontology of sustainability indicator sets. That chapter concluded that, of the two designs, the GOSIS design offers greater reuse where unseen sets of indicators need to be added to the ontology in an ad hoc fashion. In this chapter, we apply this design as a part of the development process of a recently-developed sustainability reporting software Circles of Sustainability (CoS), named after the methodology of the same name discussed by Scerri and James [2010].

By prototyping the CoS user interface, we suggest two methods of browsing and navigating sustainability indicators: Tabular view and Circular view, in which each interface is designed to assist end-users to browse heterogeneous indicator systems. We then design and conduct a user study to evaluate the overall usefulness of the CoS user interface and ease of use of the two browsing mechanisms.

The structure of the chapter is as follows. We begin by outlining the CoS user interface development in Section 5.1. We then describe the steps for evaluating the CoS user-interface with a user study in Section 5.2. Section 5.3 discusses both the quantitative and qualitative results of the user study, and finally, Section 5.4 summarises the research findings.
5.1 CoS User Interface Development

In this section, we particularly demonstrate user interface development of the CoS software to address the second research question: which mechanisms can be used to assist end-users to browse and navigate heterogeneous sustainability indicator sets used in a reporting framework? The steps in developing the CoS software – consisting of three layers: database, front-end and back-end – were discussed in Section 3.2.1.

In developing the CoS user interface, the following scenario is considered. A new user must first register (Figure 5.1). The user can then select or create a project associated with several critical issues. The user must then browse sustainability indicators from loaded indicator systems using either a Tabular view or a Circular view. Relevant indicators can then be added to critical issues. Finally, users need to add indicators and conduct an assessment of the project against the CoS. In what follows, we describe the steps of the above scenario in detail, that has been particularly adapted for the purpose of the user study evaluation.

5.1.1 Defining a Sample Project

For conducting the user study evaluation, we first create a sample project following the scenario previously described. The project considers Tehran Air Pollution.

**Tehran Air Pollution Project:**

Tehran is the hometown of the author of this thesis and is an extreme case of a community that has several long-term sustainability problems that can result in short term disasters. For example, air pollution had forced Iranian authorities to close elementary schools and kindergartens in Tehran province for 3 days in November 10, 2013\(^1\). According to the latest statistics published by the Centre for Environment and Health in Iran\(^2\), air pollution caused 4,460 deaths in 2012. This means that air pollution caused an average of seven deaths per day in this city. Tehran, the capital city of Iran, is the largest urban area in that nation and had a population of nearly 8.5 million in June 2014\(^3\). According to Naddafi et al. [2012],

\(^1\)http://www.nytimes.com/2013/01/07/world/middleeast/tehran-is-choked-by—annual—buildup—of—air—pollution.html?
\(^3\)http://worldpopulationreview.com/world-cities/tehran-population/
city ranks as one of the largest cities in Western Asia and the 19th in the world. As with other large cities, Tehran is faced with serious air quality problems. In general, 20% of the total energy of the country is consumed in Tehran. Pollutants such as PM10, SO2, NO2, HC, O3 and CO are the major air pollutants in Tehran, about 80-85% of which is produced by mobile sources of pollution.

Presenting a Series of Critical Issues

Based on preliminary research into this particular case, we pose the following issues concerning air pollution in Tehran. Using the terminology of our GOSIS ontology described in Section 4.2.2, we specify these as Critical Issues.

1. **Geographical location of the city:** Tehran’s average altitude is 1,890 metres above mean sea level. The city is located in valleys and surrounded on the north, northwest, east and southeast by high to medium high of mountain ranges of up to 3,801 meters height. These ranges block the flow of the humid wind to the city and prevent the polluted air from being carried away from the city. Thus, during winter, the lack of wind and the cold air causes the polluted air to be trapped within the city [Naddafi et al., 2012].

2. **The use of private cars:** Citizens are not encouraged to use public transport. Naddafi et al. [2012] report that the city has a capacity for 700,000 registered cars are on its streets on a daily basis. Cars account for 70% to 80% of the city’s air pollution.

3. **Poor public transport service:** Tehran Metro carries more than 2 million passengers per day. Other types of public transport, such as bus lines, are not fully developed. Furthermore, taxies are very expensive. However, this type of transport is still the most used type which also has a direct effect on the air pollution problem.

4. **Low quality of the petrol made in Iran:** As sanctions on imports have forced the country to turn to low-quality fuel, all vehicles in Tehran use the poor quality petrol. The government has tried to replace the fossil fuel consumption with the gas option

---

4http://in.reuters.com/article/2011/03/05/idINIndia-55349520110305
by encouraging drivers to convert their cars to using this alternative. However, the conversion process is costly and most drivers do not find this option economic.

5. Moving towards an industrialising city: Tehran, as Iran’s capital and one of the mega-cities in the world, is moving towards industrialisation to strengthen its economy. The actions to support this aspect of the city have polluted Tehran, both directly and indirectly. On the other hand, the severity of the air pollution at certain times of the year has forced the government to shut down the city for a few days by asking people to stay at their homes. For example, in November 2013, kindergartens, elementary schools and universities were closed for three days. This action has considerable drawbacks on the economy of Tehran and the country as a whole.

5.1.2 Browsing Mechanisms of Indicator Sets

In pursuit of supporting the user navigation of semantic heterogeneity of sustainability indicator sets, two mechanisms of browsing indicators are used for translating the back-end into the front-end interface. As discussed in Section 3.2.3, we decided to compare a text-based with a visual-based browsing approach. These are described further below.

The Tabular view is a text-based interface that presents information about indicators in a series of columns. This method provides a flow of important features including: IndicatorSet, Source, Category, Indicator, Title, Description and Subdomain(s). This view assists users to find information in a relative way and suggests an overall view of the indicator sets structure for those users who are less familiar with sustainability concepts. More importantly, this search option simplifies the process of finding relevant indicators because it can be performed using indicator names, specific keyword or particular organisations.

By contrast, the Circular view is a visual interface (see Section 3.2.3) that displays indicators in a circle divided into four domains, each of which is divided into seven sections (Subdomains). By clicking on a Domain the names of all subdomains appear at the left corner of the page and a list of related indicators and their details are illustrated at the right side of the page. In this view, reading subdomains is a good way to find relevant indicators, and is more suitable for those users who are more familiar with sustainability concepts.
<table>
<thead>
<tr>
<th>Set</th>
<th>Source</th>
<th>Category</th>
<th>Identifier</th>
<th>Title</th>
<th>Description</th>
<th>Subdomain(s)</th>
<th>Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>OECD Green Growth</td>
<td>OECD Green Growth</td>
<td>Environmental quality of life</td>
<td></td>
<td>15. Access to sewage treatment and drinking water; 15.2. Population with sustainable access to safe drinking water</td>
<td>Indicators on public access to an improved drinking water source, and to an improved sanitation facility, used to monitor the Millennium Development Goals (MDGs), are given as complements.</td>
<td>Ecology Water and Air</td>
<td>[No actions available]</td>
</tr>
<tr>
<td>UN Habitat Agenda Indicators</td>
<td>UN Habitat</td>
<td>Environmental Management</td>
<td></td>
<td>extensive indicator 10: houses in hazardous locations</td>
<td></td>
<td>Ecology Water and Air</td>
<td>[No actions available]</td>
</tr>
<tr>
<td>UN Habitat Agenda Indicators</td>
<td>UN Habitat</td>
<td>Environmental Management</td>
<td></td>
<td>check-list 5: disaster prevention and mitigation instruments</td>
<td></td>
<td>Ecology Water and Air</td>
<td>[No actions available]</td>
</tr>
<tr>
<td>UN Habitat Agenda Indicators</td>
<td>UN Habitat</td>
<td>Social development and eradication of poverty</td>
<td></td>
<td>check-list 4: gender inclusion</td>
<td></td>
<td>Ecology Water and Air</td>
<td>[No actions available]</td>
</tr>
</tbody>
</table>

*Figure 5.2: Screenshot of Tabular View*
Figure 5.3: Screenshot of Circular View
Snapshots of both views are shown in Figures 5.2 and 5.3. In a comparison of the two views, the **Tabular** view is a *component* whose cost of development is affordable for any given user interface. Whereas the **Circular** view is a *prototype* and is therefore comparatively costly to develop. In addition, this view features the CoS methodology shown in Figure 3.1, which recalls the concept of “Skeuomorphism” at user interfaces, in which a design is taken from the real world to recall the physical components. In this way, the Circular view can be regarded as a conceptual representation of sustainability indicators, whereas the Tabular view is a data structural representation.

At the next stage of the CoS user interface development, we load the GRI and OECD indicator sets by applying the GOSIS design candidate. The details of preparing the datasets and populating the ontology were given in Section 4.5.1.

### 5.2 User Interface Evaluation

In order to evaluate the user interface of the CoS software developed in the previous section, we conduct a user study. Our reasons for choosing this method as the evaluation approach are given in Section 3.2.5. The development steps of user study are described next.

#### 5.2.1 User Study Objectives

Given the second focus of this research, we deployed two mechanisms (Tabular and Circular views) for browsing heterogeneous sustainability indicator sets in Section 5.1.2. Accordingly, sets of objectives are identified for the user study from perspectives of two types of participants (*expert* and *less-expert*). (i) The first goal aims to evaluate the usefulness of the CoS software overall. (ii) The second goal compares the two methods of browsing semantic heterogeneity of sustainability indicators that are designed in the CoS user-interface.

Reviewing the literature and considering two objectives of this research – overall usefulness of the software and comparing ease of use for two ways to browse the indicators – we have decided on the *satisfaction* dimension and two specific usability measures: *perceived*
usefulness and perceived ease of use. A discussion of related work on usability measures appears in Section 3.2.4.

Based on the above objectives and two usability measures, three micro variables are identified: Expertise, Perceived usefulness of the software and Perceived ease of use of two browsing mechanisms which lead us to define the descriptive variables.

5.2.2 Descriptive Variables

Following the micro variables, twelve descriptive variables are specified, which form the design of the questions and basis analysis for them.

**Expertise**

1. EX1: Prior knowledge of sustainability concepts, theories, approaches and methods.

2. EX2: Prior knowledge of indicator systems (such as the GRI, OECD and UN).

3. EX3: Prior knowledge of sustainability reporting frameworks and specifically the CoS methodology (for example project, critical issues, project assessment).

**Overall Usefulness of the Software**

4. OU1: Satisfaction level of using software for accomplishing tasks more quickly.

5. OU2: Satisfaction level of using software in sustainability assessment for improving users’ performance, productivity and effectiveness.

6. OU3: Satisfaction level of easier use of software in sustainability assessment.

7. OU4: Satisfaction level of using the software in everyday sustainability assessment and reporting practices.

**Ease of Use for Semantic Heterogeneity of browsing indicators**

8. SH1: Satisfaction level to learn operating with both views more easily.
9. SH2: Satisfaction level to get both views to do what the user wants to do.

10. SH3: Satisfaction level of interaction with both views from clarity, comprehension and flexibility perspectives.

11. SH4: Satisfaction level in becoming skilful in using both views.

12. SH5: Satisfaction level to ease of use of both views.

<table>
<thead>
<tr>
<th>Micro Variables</th>
<th>Usability Measures</th>
<th>Descriptive Variables</th>
<th>Questions</th>
<th>Types of Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expertise</td>
<td>Background Knowledge</td>
<td>EX1, EX2, EX3</td>
<td>Q1, Q2, Q3, Q4, Q5</td>
<td>4-point Likert scale</td>
</tr>
<tr>
<td>Overall Usefulness</td>
<td>Perceived Usefulness</td>
<td>OU1, OU2, OU3, OU4</td>
<td>Q6, Q7, Q8, Q9, Q10, Q11</td>
<td>5-point Likert scale</td>
</tr>
<tr>
<td>Semantic Heterogeneity</td>
<td>Perceived Ease of Use</td>
<td>SH1, SH2, SH3, SH4, SH5</td>
<td>Q13, Q19, Q14, Q20, Q15, Q16, Q28, Q21, Q17, Q23, Q18, Q24</td>
<td>5-point Likert scale</td>
</tr>
</tbody>
</table>

Table 5.1: Mapping between Variables and Questions

### 5.2.3 Hypotheses

Prior to conducting the user study and collecting the results, we also define several hypotheses by identifying the relations between micro variables: expertise, overall usefulness of the software and semantic heterogeneity of browsing indicators, which also reflect two research objectives. The hypotheses are adapted from a study by McGrenere et al. [2002] as follows.

**Participant Hypothesis:** Participants will probably have prior knowledge of sustainability domain and the CoS software. However, there will be more “less-expert” users than “expert” users who have limited knowledge about reporting organisations.

**Usage Hypothesis:** Most participants will find the CoS software as a helpful tool to enhance, improve and accomplish their ability in sustainability assessment.
**Good Idea Hypothesis:** Two ways of browsing indicators will be easily understood and will be considered a good idea. It will be easy to switch between interfaces.

**Satisfaction Hypothesis:** Most participants will be satisfied with the CoS software overall.

**Navigation Hypothesis:** Most participants will benefit from general guidelines provided in the interface and progressively will feel that they are better able to perform the tasks.

**Learnability Hypothesis:** Using the Tabular view will provide more opportunities to learn the CoS software and its associated methodology.

**Functionality Hypothesis:** Using Circular view will facilitate most features of the CoS user interface.

Using these hypotheses help us to analyse the data collected from the participants after conducting the experiment as discussed in Section 5.3.

### 5.2.4 Questionnaire

After identifying the research objectives, variables and hypotheses of the user study, a set of questions are then developed for measuring the variables. In the current user study, we design a questionnaire that contains several closed and open questions. This allows us to collect quantitative information for particular subjects as well as gaining the insight into participants’ responses. The questions’ topic associated with variables are given in Section 5.2.2. The granularity between variables and questions is `<one:many>`, meaning that each variable may correspond to one or more than one question(s). The questionnaire is divided into the three sections. The first contains 5 closed questions with a scale of 1-4 and assesses the background knowledge of the participants about sustainability concepts and reporting
organisations. Participants are asked to answer this section before using the software and answering the other two sections. The second section of the questionnaire contains 6 questions in a 5-point Likert Scale and assesses the overall usability of the software. The third section contains two subsections for measuring “perceived ease of use” for the Tabular and Circular views respectively. In each subsection, 6 questions in the 5-point Likert Scale measure the users’ experience of using each view from different aspects of usability, namely: accomplishing tasks quickly, improving performance, increasing productivity, enhancing effectiveness, performing easy assessment and usefulness.

The questions are shown in Appendix C, Figures C.1 to C.5. Table 5.1 summarises the mapping between variables and questions. The types of the questions are also specified in the last row of the table. In addition, two open questions (Q12, Q25) and an overall feedback statement with the software are also given for collecting the qualitative responses. The time required to complete the experiment and answer the questionnaire is estimated approximately 20-30 minutes.

5.2.5 Characteristics of Target Audiences

One of the main phases of developing the user study is specifying the characteristics of the target audience. After a series of discussions with our social scientist experts of the broader project (see Section 3.2.1) about the choices of users, we selected the participants who had prior knowledge about concepts and theories concerning the sustainability domain, although the levels of their background knowledge were varied. Other factors such as gender and age were not considered in the selection criteria. In recruiting participants, we approached individuals using the snowball approach, which was based on the recommendations of other participants or researchers. Most participants were likely to have been invited through recommendations by people associated with the Global Compact Cities Programme6. These users were considered to be broadly representative of future users of the software, including members of universities, governments, community-based organisations and companies. On average our sample was, however, more aware about issues of urban sustainability than we

6http://citiesprogramme.com/
would expect of most users. As we discuss, this may skew the results favourably, but may also cause users to be more critical of defects in the software.

**Sample Size of Participants**

Following the above approach, the number of potential participants reached 38, although, 26 users submitted their responses. However, among them, only 20 answered all three sections of the questionnaire.

**5.2.6 Statistical Tests**

In choosing the appropriate statistical test, we first defined the type of the usability variables used in the questionnaire. Our variables are *ordinal*, meaning that there is a clear ordering for the variable values. For example, in the first phase of experiment the numeric responses rank from 1 to 4 depending on the participant’s familiarity with sustainability concepts.

Given the ordinal type of the user study variables and the assumption that the population (participants’ responses) are not normally distributed, we used *Spearman’s Rank Correlation Coefficient* and *Wilcoxon Signed-ranked* tests.

The Spearman’s Rank Correlation Coefficient (Spearman’s rho) test is chosen for analysing the results of the second section of the questionnaire, measuring overall usefulness of the software. It is a non-parametric test and the significant difference indicates whether there is a correlation between the two variables by ranging from $-1.00$ to $+1.00$. This coefficient indicates the degree that low or high scores on one variable tend to go with low or high scores on another variable, where 1 represents a total positive correlation, 0 is no correlation and $-1$ is a total negative correlation. We used this test to identify strong correlations at significant difference of .05 between different aspects of usefulness of the CoS software.

The Wilcoxon Signed-ranked test is chosen to analyse the results for the third section of the questionnaire. It is also a non-parametric test and is used when the data cannot be assumed to be normally distributed, in which case the use of *t-test* is inappropriate. This test compares two sets of scores from the same participants investigating any changes from one time point to another, or when individuals are subjected to more than one condition.
We used this test at significant difference of .05 to understand the preference of participants in working with two ways of browsing indicator sets (Tabular view versus Circular view).

5.2.7 Tasks

In developing the user study, participants were expected to accomplish the tasks below which are based on the scenario given in the development of the CoS user interface in Section 5.1.

1. Add relevant indicators to issues of the sample project using both user interfaces.

2. Assess the sample project and adding indicators.

Scripts

In order to guide the participants in performing these tasks, we also provided a set of instructions for using the CoS software and accomplishing the tasks within Script A and Script B that are presented in Appendix D, Figures D.1 to D.8. The scripts have similar steps with the only difference being in choosing which interface of browsing indicators to use at the beginning of the experiment. Script A asks users to choose the Tabular view first, while script B asks them to choose the Circular view first. Given the number of the participants, there is an even distribution of both scripts between them to reduce the risk that the use of one view will dominate the responses (that is 13 users received Script A and the other 13 received Script B). In other words, by alternating using two views, the participants’ responses to the third section of the questionnaire – related to comparing two views – are not biased to only one way of browsing indicators. Prior to starting the questionnaire, participants were asked to agree to the ethics of the user study, also given in Appendix B, Figures B.1 to B.4.

Next, we piloted the scripts and questionnaire with five users who had general knowledge of sustainability indicator sets and were partially familiar with the CoS software. The scripts were improved over some iterations by addressing the comments and feedback from pilot users. Finally, we conducted the user study and collected the results over a 10-day period from 8th to 18th March 2014. The participants were contacted via individual emails and received an electronic copy of the scripts (A or B) and a link to the survey of the questionnaire.
5.3 Results and Discussion

In this section, we present quantitative and qualitative data collected from the user study experiment described in previous section. The subsequent analysis of the results are also provided. Quantitative data contains the participants responses to 23 closed questions from three sections of the questionnaire and quantitative data discusses participants’ comments and feedback to two open questions.

5.3.1 Quantitative Results

Here, we present the data collected from quantitative results. IBM SPSS Statistics software (version 20)\(^7\) is used to analyse the results. Given three sections of the questionnaire, collected data is categorised into three distinct sub-categories as follow.

\(^7\)http://www-03.ibm.com/software/products/en/spss-stats-standard
Background Knowledge of the Participants

Prior to using the CoS software, participants were asked to answer the first section of the questionnaire which discusses their background knowledge about sustainability concepts, reporting organisations (e.g. GRI, OECD, UN) and CoS methodology. This section contains five Likert questions with the answers: “Very Familiar”, “Somewhat Familiar”, “A Little Familiar” and “Not Familiar”, which are ranked from 1 to 4 respectively. The responses to each question totalled 26, which indicates the fact that all users participated in this section.

Since participants’ responses provides a distribution of a single variable, for data analysis, we used descriptive statistics of the frequency of responses while each response is correspondence to a numeric value. The results are shown in the histogram in Figure 5.4, where the $x$ axis represents four types of responses in each question and the $y$ axis summarises the corresponding mean value.

Turning to the type of the responses: “Very Familiar” was chosen the least (zero) in Q3 and Q4. In contrast, “Somewhat Familiar” was the most frequent response with a mean value of 15 in Q1, which is followed by “Not Familiar” as the second most frequent response with a mean value of 13.50 in Q3 and Q4. The other type of response, “A Little Familiar” scored differently ranging from 5 to 10 of the mean value in all questions.

Looking at the questions, almost two-thirds of the users ranked their knowledge level with Q1 as “Somewhat familiar”, whereas the others ranked it as having various degrees of familiarity. Similarly, on average fewer than two-thirds of users ranked Q3 ad Q4 as “Not Familiar”, while the others had “Somewhat” or “A Little” knowledge about these questions. There were no users who had been very familiar with the aforementioned questions. Finally, the participants’ knowledge about Q2 varied in four types of the responses, where almost half of users were “A Little Familiar” or “Not Familiar”, and the other half scored this question as “Somewhat Familiar”; a few users had ranked their knowledge towards Q2 as “Very Familiar”.

These results can be interpreted in a way that most participants are familiar with sustainability concepts at a reasonable level, whereas, only a few – almost one-third of participants – had known about reporting organisations specifically. In addition, two-thirds of the par-
Results and Discussion

<table>
<thead>
<tr>
<th></th>
<th>Q6</th>
<th>Q7</th>
<th>Q8</th>
<th>Q9</th>
<th>Q10</th>
<th>Q11</th>
</tr>
</thead>
<tbody>
<tr>
<td>N Valid</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>N Missing</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Mean</td>
<td>1.800</td>
<td>1.400</td>
<td>1.900</td>
<td>1.650</td>
<td>1.600</td>
<td>1.700</td>
</tr>
<tr>
<td>Median</td>
<td>2.00</td>
<td>1.00</td>
<td>2.00</td>
<td>2.00</td>
<td>1.00</td>
<td>1.500</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>.767</td>
<td>.502</td>
<td>.718</td>
<td>.587</td>
<td>.820</td>
<td>.864</td>
</tr>
<tr>
<td>Range</td>
<td>3.00</td>
<td>1.00</td>
<td>2.00</td>
<td>2.00</td>
<td>3.00</td>
<td>3.00</td>
</tr>
<tr>
<td>Minimum</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Maximum</td>
<td>4.00</td>
<td>2.00</td>
<td>3.00</td>
<td>3.00</td>
<td>4.00</td>
<td>4.00</td>
</tr>
</tbody>
</table>

Table 5.2: Descriptive Statistics of Six Questions - Perceived Usefulness

Participants had some familiarity with CoS methodology.

These findings prove our participant hypothesis (Section 5.2.3) that most participants satisfy the condition of familiarity with sustainability domain and the CoS methodology. However, only a few experts who know about sustainability indicators sets. It can be concluded that the user study experiment is conducted for the proper sample size of participants who have some prior knowledge about the domain. Therefore, their responses are accurate and their feedback can lead us to a more reliable analysis of the data.

Overall Usefulness

The second section of the questionnaire measured perceived usefulness of the CoS software. This section had six questions with five response types: “Extremely Likely”, “Slightly Likely”, “Neither”, “Slightly Unlikely” and “Extremely Unlikely”, which ranked from 1 to 5 respectively. Lower values represent higher satisfactions of the participants with regards to the overall usability of the software. As discussed in Section 5.2.2, the questions’ topics are correspond to the usability variables adapted from the study by McGrenere et al. [2002] which are discussed in Section 5.2.2. We conducted the Spearman’s rho test to identify significant correlations between the usability variables (given in Section 5.2.2) whose the questions are designed based on.

Tables 5.2 and 5.3 illustrate the descriptive and correlation results in a tabular form. The Null Hypothesis exists when there is no significant correlation and alternative hypothesis is
valid when there is a strong correlation. The sample size is shown by $N$, which is 20 in our user study indicating that only 20 participants out of 26 completed this section. The results indicate that Q7 shows strong correlations with other five questions. In particular, the highest correlation value is between (Q7,Q10) of 0.895, which is significant at the confidence level of .01. The second highest correlation values are seen between pairs of (Q6,Q11) and (Q7,Q11) of 0.745 and 0.747 respectively, which are also significant at the confidence level of .01. In contrast, the weakest correlation exists between pairs of (Q9,Q11) and (Q6,Q9) with the values of 0.369 and 0.486 respectively.

These figures support some of our hypotheses given in Section 5.2.3 but reject others. Q7 is about improving participants’ performance using the CoS software and it has strong

<table>
<thead>
<tr>
<th></th>
<th>Q6</th>
<th>Q7</th>
<th>Q8</th>
<th>Q9</th>
<th>Q10</th>
<th>Q11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation Coefficient</td>
<td>1.000</td>
<td>.656**</td>
<td>.515*</td>
<td>.486*</td>
<td>.715**</td>
<td>.745**</td>
</tr>
<tr>
<td>Q6</td>
<td>Sig. (2-tailed)</td>
<td>–</td>
<td>.002</td>
<td>.020</td>
<td>.030</td>
<td>.000</td>
</tr>
<tr>
<td>N</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td>.656**</td>
<td>1.000</td>
<td>.694**</td>
<td>.685**</td>
<td>.895**</td>
<td>.747**</td>
</tr>
<tr>
<td>Q7</td>
<td>Sig. (2-tailed)</td>
<td>.002</td>
<td>–</td>
<td>.001</td>
<td>.001</td>
<td>.000</td>
</tr>
<tr>
<td>N</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td>.515*</td>
<td>.694**</td>
<td>1.000</td>
<td>.657**</td>
<td>.549*</td>
<td>.477*</td>
</tr>
<tr>
<td>Q8</td>
<td>Sig. (2-tailed)</td>
<td>.020</td>
<td>.001</td>
<td>–</td>
<td>.002</td>
<td>.012</td>
</tr>
<tr>
<td>N</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td>.486*</td>
<td>.685**</td>
<td>.657**</td>
<td>1.000</td>
<td>.710**</td>
<td>.369</td>
</tr>
<tr>
<td>Q9</td>
<td>Sig. (2-tailed)</td>
<td>.030</td>
<td>.001</td>
<td>.002</td>
<td>–</td>
<td>.000</td>
</tr>
<tr>
<td>N</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td>.715**</td>
<td>.895**</td>
<td>.549*</td>
<td>.710**</td>
<td>1.000</td>
<td>.705**</td>
</tr>
<tr>
<td>Q10</td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.000</td>
<td>.012</td>
<td>.000</td>
<td>–</td>
</tr>
<tr>
<td>N</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td>.745**</td>
<td>.747**</td>
<td>.477*</td>
<td>.369</td>
<td>.705**</td>
<td>1.000</td>
</tr>
<tr>
<td>Q11</td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.000</td>
<td>.034</td>
<td>.109</td>
<td>.001</td>
</tr>
<tr>
<td>N</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

*Correlation is significant at the 0.01 level (2-tailed)
**Correlation is significant at the 0.05 level (2-tailed)
Results and Discussion

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q13</td>
<td>19</td>
<td>2.47</td>
<td>1.35</td>
<td>1.00</td>
<td>5.00</td>
</tr>
<tr>
<td>Q14</td>
<td>19</td>
<td>2.63</td>
<td>1.06</td>
<td>1.00</td>
<td>5.00</td>
</tr>
<tr>
<td>Q15</td>
<td>19</td>
<td>2.74</td>
<td>1.19</td>
<td>1.00</td>
<td>5.00</td>
</tr>
<tr>
<td>Q16</td>
<td>19</td>
<td>2.90</td>
<td>1.15</td>
<td>1.00</td>
<td>5.00</td>
</tr>
<tr>
<td>Q17</td>
<td>19</td>
<td>2.26</td>
<td>1.15</td>
<td>1.00</td>
<td>5.00</td>
</tr>
<tr>
<td>Q18</td>
<td>19</td>
<td>2.32</td>
<td>1.21</td>
<td>1.00</td>
<td>5.00</td>
</tr>
<tr>
<td>Q19</td>
<td>18</td>
<td>1.72</td>
<td>.83</td>
<td>1.00</td>
<td>4.00</td>
</tr>
<tr>
<td>Q20</td>
<td>19</td>
<td>1.80</td>
<td>.79</td>
<td>1.00</td>
<td>3.00</td>
</tr>
<tr>
<td>Q21</td>
<td>19</td>
<td>1.79</td>
<td>.63</td>
<td>1.00</td>
<td>3.00</td>
</tr>
<tr>
<td>Q22</td>
<td>19</td>
<td>1.79</td>
<td>.85</td>
<td>1.00</td>
<td>4.00</td>
</tr>
<tr>
<td>Q23</td>
<td>19</td>
<td>1.68</td>
<td>.88</td>
<td>1.00</td>
<td>4.00</td>
</tr>
<tr>
<td>Q24</td>
<td>19</td>
<td>2.32</td>
<td>1.20</td>
<td>1.00</td>
<td>5.00</td>
</tr>
</tbody>
</table>

Table 5.4: Signed Wilcoxon Descriptive Statistics

correlations with all other questions, which focus on individual aspects of the usability of the software. This supports the satisfaction hypothesis that says most participants will be satisfied with the CoS software overall. Furthermore, Q11 suggests using the CoS software is a useful tool at sustainability assessment and has strong correlations with two other questions, which break down the usability aspects into two sub-tasks namely: Accomplishing tasks more quickly (Q6), and improving performance (Q7). This partially supports the usage hypothesis that the CoS software will improve the participants’ ability at sustainability assessment. However, the weak correlations between pairs of (Q9,Q11) and (Q6,Q9) reject the hypothesis that using CoS will enhance the effectiveness of sustainability assessment, which is the content of Q9. The reason could be the “effectiveness of the software” was not clear for all participants, who ranked it very low. We think the reasons for this are related to the technical limitations of the current implementation of the software which are listed in Section 5.3.2.

Ease of Use of Browsing Indicator Sets

The third section of the questionnaire aimed at comparing users’ satisfaction with the perceived ease of use of two ways of browsing indicators: Tabular view versus Circular view. Twelve 5-point Likert scale questions are paired for each view using the variables discussed
Results and Discussion

Table 5.5: Signed Wilcoxon Test - Significant Difference and Z-value - Perceived Ease of Use

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean Rank</th>
<th>Sum of Ranks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q13-Q19</td>
<td>Negative Ranks</td>
<td>7&lt;sup&gt;a&lt;/sup&gt;</td>
<td>7.50</td>
</tr>
<tr>
<td></td>
<td>Positive Ranks</td>
<td>4&lt;sup&gt;b&lt;/sup&gt;</td>
<td>3.38</td>
</tr>
<tr>
<td></td>
<td>Ties</td>
<td>7&lt;sup&gt;c&lt;/sup&gt;</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>18</td>
<td>–</td>
</tr>
<tr>
<td>Q14-Q20</td>
<td>Negative Ranks</td>
<td>10&lt;sup&gt;d&lt;/sup&gt;</td>
<td>6.90</td>
</tr>
<tr>
<td></td>
<td>Positive Ranks</td>
<td>2&lt;sup&gt;e&lt;/sup&gt;</td>
<td>4.50</td>
</tr>
<tr>
<td></td>
<td>Ties</td>
<td>7&lt;sup&gt;f&lt;/sup&gt;</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>19</td>
<td>–</td>
</tr>
<tr>
<td>Q15-Q21</td>
<td>Negative Ranks</td>
<td>11&lt;sup&gt;g&lt;/sup&gt;</td>
<td>6.41</td>
</tr>
<tr>
<td></td>
<td>Positive Ranks</td>
<td>1&lt;sup&gt;h&lt;/sup&gt;</td>
<td>7.50</td>
</tr>
<tr>
<td></td>
<td>Ties</td>
<td>7&lt;sup&gt;i&lt;/sup&gt;</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>19</td>
<td>–</td>
</tr>
<tr>
<td>Q16-Q22</td>
<td>Negative Ranks</td>
<td>13&lt;sup&gt;j&lt;/sup&gt;</td>
<td>7.35</td>
</tr>
<tr>
<td></td>
<td>Positive Ranks</td>
<td>1&lt;sup&gt;k&lt;/sup&gt;</td>
<td>9.50</td>
</tr>
<tr>
<td></td>
<td>Ties</td>
<td>5&lt;sup&gt;l&lt;/sup&gt;</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>19</td>
<td>–</td>
</tr>
<tr>
<td>Q17-Q23</td>
<td>Negative Ranks</td>
<td>8&lt;sup&gt;m&lt;/sup&gt;</td>
<td>5.19</td>
</tr>
<tr>
<td></td>
<td>Positive Ranks</td>
<td>1&lt;sup&gt;n&lt;/sup&gt;</td>
<td>3.50</td>
</tr>
<tr>
<td></td>
<td>Ties</td>
<td>10&lt;sup&gt;o&lt;/sup&gt;</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>19</td>
<td>–</td>
</tr>
<tr>
<td>Q18-Q24</td>
<td>Negative Ranks</td>
<td>2&lt;sup&gt;p&lt;/sup&gt;</td>
<td>2.50</td>
</tr>
<tr>
<td></td>
<td>Positive Ranks</td>
<td>2&lt;sup&gt;q&lt;/sup&gt;</td>
<td>2.50</td>
</tr>
<tr>
<td></td>
<td>Ties</td>
<td>15&lt;sup&gt;r&lt;/sup&gt;</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>19</td>
<td>–</td>
</tr>
</tbody>
</table>

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>Q19&lt;Q13</td>
<td>b.</td>
<td>Q19&gt;Q13</td>
</tr>
<tr>
<td>d.</td>
<td>Q20&lt;Q14</td>
<td>e.</td>
<td>Q20&gt;Q14</td>
</tr>
<tr>
<td>g.</td>
<td>Q21&lt;Q15</td>
<td>h.</td>
<td>Q21&gt;Q15</td>
</tr>
<tr>
<td>j.</td>
<td>Q22&lt;Q16</td>
<td>k.</td>
<td>Q22&gt;Q16</td>
</tr>
<tr>
<td>m.</td>
<td>Q23&lt;Q17</td>
<td>n.</td>
<td>Q23&gt;Q17</td>
</tr>
<tr>
<td>p.</td>
<td>Q24&lt;Q18</td>
<td>q.</td>
<td>Q24&gt;Q18</td>
</tr>
</tbody>
</table>

in Section 5.2.2. The questions in this section also have the same type of the responses from previous section. For analysing the results, we used the Signed Wilcoxon test for the paired responses, for example Q13 (Tabular view) versus Q19 (Circular view).

Tables 5.4 and 5.5 feature descriptive and rank results of the Wilcoxon test. The results suggest that questions related to the Circular view rank lower than Tabular view, in which, lower values indicate greater satisfaction. As a result, four pairs out of six questions – (Q14,Q20), (Q15,Q21), (Q16,Q22) and (Q17,Q23) shown in Table 5.6 – ranked positively at significant difference of .05. This indicates that users were more satisfied working with the
Table 5.6: Signed Wilcoxon Test - Perceived Ease of Use

<table>
<thead>
<tr>
<th></th>
<th>Q13-Q19</th>
<th>Q14-Q20</th>
<th>Q15-Q21</th>
<th>Q16-Q22</th>
<th>Q17-Q23</th>
<th>Q18-Q24</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z</td>
<td>-1.749&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-2.434&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-2.516&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-2.745&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-2.326&lt;sup&gt;b&lt;/sup&gt;</td>
<td>.000&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Asymp. Sig. (2-tailed)</td>
<td>.080</td>
<td>.015</td>
<td>.012</td>
<td>.006</td>
<td>.020</td>
<td>1.000</td>
</tr>
</tbody>
</table>

a. Wilcoxon Signed Ranks Test  
b. Based on Positive Ranks  
c. The Sum of Negative Ranks Equals the Sum of Positive Ranks

CoS features using the Circular view rather than the Tabular view. Those features in which the Circular view is more preferred are: (i) Easy to understand what to do, (ii) Interaction is clearer and understandable, (iii) It is more flexible to interact with and (iv) It is easy to become skilful in its use.

Further analyses of the results indicate several findings. First, the Circular view is the preferred interface for browsing heterogeneous indicators by most participants. This confirms the functionality hypothesis, given in Section 5.2.3, which is using Circular view facilitates most features of the CoS software. Second, the learnability hypothesis of the Tabular view is not supported because the Wilcoxon signed test did not find any significant difference between the scores of the two views for the question of “Learning to operate with which view is more easier?”. These findings conclude that the question, “Which view assists users more to use CoS software?” does not have a definite answer. It depends on which feature of the CoS software are considered. Although, the Circular view is preferable for most reporting procedures, the Tabular view is also ranked higher for some of the tasks. Subsequent analyses of qualitative participants’ feedback clarify these points.
<table>
<thead>
<tr>
<th>Tabular view</th>
<th>Circular view</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The Tabular view is easier to browse and compare among indicator sets.</td>
<td>1. It provides a comparative visual experience of the assessment process.</td>
</tr>
<tr>
<td>2. The Tabular view is better for specific searching.</td>
<td>2. The Circular view is better for a graphic display which then could be used for reporting more easily, quicker and more pleasant way.</td>
</tr>
<tr>
<td>3. The Tabular view is less visually oriented and more complex to use.</td>
<td>3. It is visual and communicates to the user more easily and it is visually easier to navigate.</td>
</tr>
<tr>
<td>4. For complicated models with numerous indicators, the Tabular view would be more efficient and usable.</td>
<td>4. It is easier to see several different indicators simultaneously. This is important for comparison of indicators as some are very similar.</td>
</tr>
<tr>
<td>5. The Tabular view is more preferred because there are many text related to each indicator and would be useful to view them all together at one.</td>
<td>5. The Circular view appears to assist general indicator browsing and give short details. Also indicators can be viewed according to the individual frameworks e.g. OECD.</td>
</tr>
<tr>
<td>6. The Tabular view as it represents the environment in a more holistic and realistic manner.</td>
<td>6. The Circular view is far easier to use because it only requires the user’s intuition to figure out. Each piece of the circle can easily be identified and manipulated.</td>
</tr>
<tr>
<td>7. In Tabular view it would be helpful if indicators could be ‘refined’ and sorted based on criteria such as ‘tags’, keywords, themes etc.</td>
<td>7. It would be useful to have the option to select additional areas of information per each section of the circle by simply clicking with the mouse on a single segment of the circle.</td>
</tr>
<tr>
<td>8. In the Tabular view browsing is slow, and a keyword search did not throw up the kind of indicators.</td>
<td>8. The Circular view may be more appealing for users with a language barrier.</td>
</tr>
<tr>
<td>9. There was too much text involved with the Tabular view, which made it difficult to obtain an overall picture of the different indicators.</td>
<td>9. The Circular view works well as long as the indicator can easily be categories against the four parameters.</td>
</tr>
</tbody>
</table>

*Table 5.7: Users’ Insight of Working with Two Views*
Which View is Preferable?

Q25. The software ultimately aims to help an organisation to select indicators for measuring a series of critical issues that a stakeholder group has identified.

In addressing this goal, which view is preferable?

This question investigates the insight of users’ experiences of working with the two ways of browsing indicator sets at the end of third section of the questionnaire. In Table 5.7, nine features are listed for each view extracted from users’ responses. While it is not appropriate to compare individual responses, some conclusions can be made after reviewing the responses from both views.

The Tabular view is preferred for its simplicity in presenting indicators in a table format, having a search functionality, and being less visually oriented. However, this view fails in providing users with information about the structure of indicator sets, and displays too much text at a time, which is not visually pleasant. On the other hand, the Circular view is preferred by most users as providing several advantages including: comparative visual experience of assessment process, better graphical display, being easier to navigate and communicate visually and providing short details about indicators with distinguishing their domains and sub-domains. The downside with this view is: indicators must be correlated to at least one of the Circles of Sustainability.

5.3.2 Qualitative Results

Two open questions at the end of second and third sections of the questionnaire aim at understanding participants’ insight. Some of the participants’ responses are given in Appendix E, Figures E.1 to E.4. Here we highlight the key findings as follows.

Q12. Can you imagine using the software in everyday sustainability assessment and reporting practices? What could be done or added to make it more useful?

The variety of responses to this question, lead us to divide them into three categories: Software strengths, software limitations and a suggestion list for improvements.
Software Strengths

In the course of qualitative analysis of the participants’ responses, several software strengths are found.

- The software can be useful in comparing different datasets with a sustainability reporting framework and it can be improved by a guidance of methodology steps to assist the navigation.

- The simplicity with which the user is able to navigate the site makes this an ideal tool for everyday use.

- The buttons and text clearly guide the user through each step.

- While the instructions are useful, it would not take a user long to learn the software.

- It allows the user to correct and update information.

- The software has the potential to become an important component of urban development efforts.

- The CoS software would definitely assist and support the assessment and reporting practices and process as it gives an immediate, tangible and readable results through the assessments.

Software Limitations

Our participants also highlighted some of the software limitations as follows:

- The software is a little cumbersome to navigate and uncomfortable to use. The guidance through the steps is not intuitive, which means practice and time is required to understand the steps.

- User interface is not intuitive for untrained users.

- The first task in the experiment (that is adding issues) is confusing.
Results and Discussion

• It is unclear what sort of assessment values (second task) to enter in the final screen: there is a great deal of variation between the indicators. For example, some are measured as percentages (%), others in volume (e.g. emissions particles), still others as scores. This makes it difficult to use a single system of reporting or assessment.

• The method of selecting the indicators is still unwieldy. Indicator sets are difficult to quantify and requires more description. A large part of the problem is that there are not enough relevant indicators in the system.

• Several technical issues were also detected by users:

  1. Having a delete button prominently displayed at the entry point can be risky.

  2. The automatic process of subdomain selecting by keywords is inappropriate. For example, some of the subdomains would be automatically chosen when the user writes the terms of the issue. In addition, while a sub-domain is selected if another sub-domain is also chosen, selection changes but the selection of the former sub-domain (red part) remains fixed.

  3. The system does not seem to save consistently.

  4. The “Add a new issue” button needs to be separate from the “Add a new indicator”.

  5. Unless a page is accessed through the “Add existing indicator” path, no actions are available from either Tabular view or Circular view.

Given the issues above that were highlighted from participants’ responses, it can be seen that overall the software is considered to be a useful tool for reporting practices and assessments, although there are conflicting ideas about the simplicity of using software. While some expert users benefit from the present scenario – defining a project and issues, adding indicators to issues and assessing the project against circles of sustainability – less expert users find this scenario confusing and complex. In addition, users with little knowledge about indicator sets seem confused about the adequateness of information provided in the interface. Furthermore, unclear navigation steps and less intuitive guidance are considered to be other shortcomings
of the interface. This rejects the navigation hypothesis given in Section 5.2.3. Finally, users detected several technical issues concerning the current implementation.

**Suggestions for Improvements**

Below, we present a list of improvements for the CoS software and its user-interface which are stated by our participants. We made no changes to their comments.

- Have a more visually appealing graphic design of the software.
- It would be useful if software could generate assessment “tools”. (for example ratings scales, survey questions etc.).
- A searchable “tag” system would make it easier to find relevant indicators for an assessment.
- A search function should be added to the Circular view.
- Have a structured framework against which a user can perform each assessment. This would help to add more functionality to the “Browse Indicators” page. The search function seems to work well, but it would be great if adding indicators to issues from this table is enabled.
- Assessing indicators is clearer than adding issues and the charting tool looks really useful and user-friendly. This is mainly because the software instructions provided could be more clear. The instructions might also benefit from some more context and explanations regarding the purpose of adding issue.
- It would be better and save a lot of time if the system was fully populated with dozens of relevant indicators in each subdomain.
- Possibly the process may be facilitated and make it more user-friendly through the dashboard selection.
- The circles inside the process would be better having the subdomains named directly against them or coming up with spatial contiguity with the cursor.
• Based on the level of knowledge, defining different sustainability indicators would be useful.

• The indicators can be detailed for the experts, but the indicators of the same assessment should be designed more simpler for non-professional people.

5.4 Summary

The overall objective of the present research is developing knowledge bases utilising ontologies to tackle heterogeneity issues of data representation for two related domains of sustainability indicator sets and disaster management. In this chapter, we addressed the second focus, which examined how Semantic Web mechanisms can help end-users to browse heterogeneous indicators to develop a sustainability reporting framework. Through the development of the CoS software, we applied the generic ontology design to model the indicator systems and suggested two mechanisms for browsing indicator sets. We conducted a user study consisting a questionnaire of 25 questions to analyse the participants’ satisfaction with using the CoS software and user interface from two usability measures of perceived usefulness and perceived ease of use. The questionnaire was divided into three sections including: background knowledge (5 closed questions), overall usefulness of the software (6 closed and 1 open question) and ease of use of browsing sustainability indicator sets (12 closed and 1 open questions). Accordingly, we analysed the quantitative and qualitative results following the three perspectives.

First, the results of the background knowledge suggest that most participants had some familiarity with sustainability domain and the CoS methodology. However, only a few experts who knew about sustainability indicators sets. This concluded that the user study was conducted for a proper sample size of participants and their feedback was reliable and precise. Second, the subsequent analyses of quantitative results of the second section of questionnaire support the satisfaction hypothesis; that is, most participants were satisfied using the software. Another conclusion is the proof of usage hypothesis: the CoS framework assisted users to accomplish tasks more quickly, increased their productivity and enhanced the effectiveness of sustainability assessment with the higher level of satisfaction. Further, a
review of users’ feedback suggests that the CoS software is generally considered to be a useful tool for reporting practices and assessments, although there are conflicting ideas about the simplicity of using software from expert and less-expert users points of view concerning the simplicity of using software. In addition, inadequate information about sustainability indicators provided in the interface and unclear navigation steps and less intuitive guidance are seen to be other shortcomings. This conclusion therefore rejects the navigation hypothesis, which is not surprising for a prototype system that aims to support a complex set of features.

The results analysis of the third section of the questionnaire indicates several key findings. (i) The functionality hypothesis is validated through the Circular view, which is the preferred interface to browse semantic heterogeneity of indicators by most participants. (ii) Learnability hypothesis about the Tabular view is not supported. (iii) The answer to the question, “Which view assists users more to use CoS software?” does not have a definite answer. It is a relative answer, depending on the functions of the CoS software. Although the Circular view is preferred for most reporting procedures, the Tabular view is also considered to be a better alternative for some features. This suggests that the time and effort spent developing an alternative view and navigation device needs to be balanced by the profile of the users and the tasks they are looking to perform. For comparing indicators, the Tabular view is preferred. For exploring indicator types, and assigning values to them, the Circular view is preferred.
Chapter 6

Geotagging Tweets in the Context of Disaster Management

There are in fact two things, science and opinion; the former begets knowledge, the latter ignorance.
–Hippocrates

Following on from discussions about ontology design and user interface navigation for sustainability indicator sets, in this chapter we examine the problem of dealing with semantic heterogeneity with real-time semi-structured data generated by social media in the context of disaster management. While vast amount of data are generated via social media in transmitting information at the time of a disaster, evidence shows that analysis of tweets for past or present events increases the potential to detect a disastrous event and as a result, improves the emergency situation awareness [McCarthy and Boyd, 2005; Mills et al., 2009; Starbird et al., 2010; Vieweg et al., 2010; Yin et al., 2012]. A review of related work in analysing social media content is provided in Section 2.10.

In particular, Twitter plays an important role in this context because emergency tweets are valuable resources when tagged with their location information which could result in detecting unexpected events. Twitter messages contain different types of locations such as a place name, an incomplete address or a city or suburb name. One of the main challenges
in geotagging tweets is the *ambiguity* involved in the names of the location, that are called *toponyms*.

In this chapter, we treat the ambiguity of toponyms in emergency tweets’ content as another type of semantic heterogeneity issue, which is the third focus of this research in exploring semantic techniques to solve the heterogeneity issues of georeferencing tweets in the context of disaster management. To pursue this goal, we develop, OzCT, a geotagger for the given framework of OzCrisisTracker that uses *Toponym Recognition* and *Toponym Resolution* approaches to detect and resolve the heterogeneity issues of georeferences within the tweets’ content. OzCT geotagger is then evaluated against two other methods.

The structure of the chapter is organised as follows. We begin by outlining the data collection procedure in Section 6.1. Next, the two approaches used in the development of OzCT geotagger – Tweet Toponym Recognition and Tweet Toponym Resolution – are described in Sections 6.2 and 6.3. Next, the evaluation approach and measures are described in Section 6.4. This is followed by presenting and discussing the results in Section 6.5. As a further development of our model, the ontological model for toponym resolution is briefly described in Section 6.5.2. We finally summarise the key findings in Sections 6.6.

### 6.1 Data Collection

The data used in this chapter, was collected by the OzCrisisTracker Web-application. As mentioned in Section 3.3.1, OzCrisisTracker is an adaptation of the CrisisTracker open source platform that aims at improving real-time situation awareness by combining some automated clustering approaches with a crowdsourcing technique of tweets’ content. According to Rogstadius et al. [2013], a system administrator first collects the tweets through the streaming API of Twitter\(^1\) and filters the tweets by user accounts and geographic bounding boxes. One of the filters specifically applied for is that geographic bounding boxes is limited to Australia region only. Incoming tweets are then compared with previously collected tweets using a bag of words approach that detects textual similarities by a cosine similarity metric.

\(^1\)https://dev.twitter.com/docs/streaming-apis
for grouping messages. Next, the clustering process uses a probabilistic technique applying hash functions to detect near duplicates in a stream of feature vectors that constructs clusters which are also referred to as stories. Finally, with the interaction of volunteers, tweets are grouped into categories that are indications of the type of disaster (such as Fire or Flood) or relevant to the disastrous event (such as Seek help or Warning). 14 categories are distinguished through the collected tweets during the Australian Bushfire and Flood season from January to April 2013, that are listed in Table 6.1. The total number of tweets collected were nearly one hundred thousand including retweets (RT). For the experiments of this research, we filtered retweets since they contain the same geolocation as the original tweets. We also divided the whole dataset into training and testing sets based on their date of collection. The datasets are described below.

**Training Set:** The training dataset is collected from January and February 2012 (over 60,000); some samples are shown in Table 6.2. This dataset assists us to identify the heuristic rules illustrated in Table 6.2.
**Testing Set:** The testing dataset is 500 tweets sampled out of over 22,000 tweets collected during March and April 2013. The tweet proportions of each category in the testing dataset are displayed in Table 6.1. The somewhat arbitrary number of tweets randomly selected from each category was chosen based on two criteria: (i) The prevalence of the category in the training set, and (ii) The relevance of the categories to the geotagging discovered from other resources. The testing dataset is used for the evaluation process (see Section 6.4).

As discussed in Section 3.3.2, different researchers [Yin et al., 2012; Lee et al., 2013; Amitay et al., 2004] have addressed georeference identification for Web resources. Our approach uses a technique similar to Lieberman et al. [2010]. A comparison between this method and OzCT is given in the same section. Figure 6.1 illustrates the structure of our approach which consists of two main phases: tweet toponym recognition and tweet toponym resolution.

![Figure 6.1: OzCT Geotagger Structure](image-url)
6.2 Tweet Toponym Recognition

During the phase of tweet toponym recognition all geo/non-geo entities in the tweet content are identified. As shown in Figure 6.1 in trapezoid shape, this phase contains four steps. The first step is detecting noise which removes stop words and cleans tweet content from irrelevant words. The second step includes identifying geo/non-geo entities in the content which uses three methods: (i) A list of Australian road types (such as complete names and abbreviations such as “Rd” and “Street”) and specific locations (such as “airport” and “train station”) are applied to identify location indicators if they exist. (ii) Stanford Named Entity Recognition (NER) and Part of Speech Tagging (PoS) is applied to identify other possible location entities which do not contain any particular indicator. (iii) If no location entity is identified during the previous methods, the tweet content is then parsed and each lexicon is passed to the gazetteers. The reason behind this is that the low precision in statistical methods used in the Stanford NER and PoS often causes some location entities to be not recognised or to be identified as other named entities such as an Organisation or a Person. At the third step, geographic entities are passed to the geographic knowledge bases (gazetteers)\(^2\) for querying the corresponding geographical location. The details of geographic knowledge bases are given in Section 3.3.3.

To recap, algorithm 4.1 simplifies the development process of OzCT. Using this algorithm, 500 tweets are sampled from testing set, they are automatically geotagged by this algorithm one by one \((t)\). A list of location phrases \((L)\) is identified by NER and if the size is not zero, the list is then searched: first in the map of the program and, second in the geographic knowledge bases to identify geographic references. We apply this mechanism to reduce the search time and increase the overall geotagging time. Any new detected location phrase with the geographical references is then stored in the map for the next similar lookup. The complexity of lookups in the map or gazetteers is proportional to the number of the locations identified in the tweet by the NER.

\(^2\)In this chapter, we refer to gazetteers as geographic knowledge bases.
**Algorithm 6.2.1: Tweet Toponym Recognition**

```plaintext
FUNCTION OzCT geotagger(t)

Find location keywords in the content
$L ←$ List of location keywords
if $|L| = 0$
then $f ←$ No-Location
   if L exists in the map
      then $G ←$ Stored results
      else Find references in gazetteers
       $G ←$ Search results
   else if $|G| = 0$
      then $f ←$ No-Location
      else if $|G| = 1$
         then $f ←$ Definite
      else $f ←$ Ambiguous

return $(f, G)$
```

The output of the tweet toponym recognition algorithm is a tweet with one of the three statuses ($f$): *definite*, *ambiguous* and *no-location* and a collection of geographical references ($G$) retrieved from map or gazetteers, which are then applied on the tweet toponym resolution phase to resolve the ambiguities if possible.

### 6.3 Tweet Toponym Resolution

During the phase of toponym resolution, each toponym is assigned to the correct geographical coordinates. In Figure 6.1, the steps are shown in rectangle shape. This phase becomes
### Scenario Heuristic Rule Examples OzCT Geotagger Output

1. There is no adequate information about a location such as state, suburb or city.
   - No Heuristic rule can be applied
   - 17:34 The Bluff Road St Leonards VIC
   - 17:34 Bluff Road Erigolia NSW
   - 17:34 Bluff Road Devonport TAS
   - 17:34 Bluff Road Melbourne VIC
   - 17:34 Bluff Road Cedar Vale QLD
   - 17:34 Bluff Road Minlaton SA

2. The tweet content contains geo-anchor toponyms that adds extra information for the resolution phase.
   - An additional feature is added to the final address.
   - 10 kilometres south east of Edenhope, Advice, 17/02/13 12:19 PM
   - South East - Edenhope VIC

3. The tweet content has some ambiguous location references but it also contains other indicators that resolve the disambiguation.
   - By analysing the hashtag, the geotagger resolves the disambiguation between a list of retrieved location
   - 23:47 Scott St Severe Thunderstorm Warning for Heavy Rainfall and Damaging Winds #Cavendish
   - Scott Street Cavendish VIC

4. More than one geographical location is detected in the content.
   - By providing a weight vector, the geotagger assign more weight to the state and suburb which is common between geographical location results.
   - Thompons Rd near Narre Warren Rd
   - Thompons Road Newborough VIC
   - Thompons Road Watchem VIC
   - Thompons Road Balliang VIC
   - Thompons Road Dingee VIC
   - Narre Warren Road Melbourne VIC

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Heuristic Rule</th>
<th>Examples</th>
<th>OzCT Geotagger Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>No Heuristic</td>
<td>17:34 The Bluff Road St Leonards VIC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>rule can</td>
<td>17:34 Bluff Road Erigolia NSW</td>
<td></td>
</tr>
<tr>
<td></td>
<td>be applied</td>
<td>17:34 Bluff Road Devonport TAS</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>17:34 Bluff Road Melbourne VIC</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>17:34 Bluff Road Cedar Vale QLD</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>17:34 Bluff Road Minlaton SA</td>
<td></td>
</tr>
</tbody>
</table>

### Table 6.2: A Set of Heuristic Developed from Training Dataset Used in OzCT Toponym Resolution
more critical when a list of ambiguous locations is identified during the toponym recognition phase. The ultimate goal of the OzCT is disambiguating the results and detecting the definite location if possible. The root of the ambiguity can be classified into four different scenarios which lead us to define and apply *Heuristic Rules* to resolve them. The scenarios are presented in Table 6.2 and a brief description of each follows.

1. Scenario 1 refers to those tweets with no geospatial clues for further analysis. They contain some geographical references at the *street* or *suburb* levels which result in a list of ambiguous locations captured from gazetteers, but such information is not adequate to detect definite location. At this stage, no heuristic rules can be applied.

2. Unlike the first type of tweets, scenario 2 outlines another type of tweets that include some geo-anchor toponyms such as *south east*, *north east* and *southern* which lead us to apply a heuristic rule for additional features of location entity of such tweets. These representation can be used for more detail analysis of tweets.

3. Third scenario considers those Twitter messages containing other indicators that can resolve the disambiguation of the location. *Hashtags* are one of these indicators that in the majority of cases refer to *city* or *state* of the particular geographic reference appearing in the content. Analysing and applying hashtags lead us to resolve almost half of the ambiguous location results.

4. The final type of tweets are the ones with more than one location entity. Once all the possible location references are identified in the toponym recognition phase, the heuristic rule is defined to calculate a *weight vector* which assign more weight to those suburbs, cities and states in common between the location candidates. Maximum weight(s) then resolves the ambiguity if possible.
<table>
<thead>
<tr>
<th>Ontology/LOD Resource Used</th>
<th>Target Domain</th>
<th>Corpora Used</th>
<th>Dataset Evaluated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ritter [Ritter et al., 2011]</td>
<td>Freebase</td>
<td>Open Domain</td>
<td>6M Tweets</td>
</tr>
<tr>
<td>Laniado and Mika [Laniado and Mika, 2010]</td>
<td>Freebase</td>
<td>Open Domain</td>
<td>539432689 Tweets</td>
</tr>
<tr>
<td>Rowe [Rowe and Stankovic, 2012]</td>
<td>DBpedia</td>
<td>Conferences</td>
<td>1082 Tweets</td>
</tr>
<tr>
<td>OzCT Geotagger</td>
<td>SIOC, SIOCT, GN</td>
<td>OzCrisisTracker Application</td>
<td>900000 Tweets</td>
</tr>
</tbody>
</table>

Table 6.3: Ontology-Based Semantic Annotation Adapted [Bontcheva and Rout, 2012] and Compared with the OzCT Geotagger

<table>
<thead>
<tr>
<th>OzCT Geotagger</th>
<th>Ground Truth - Human Judgement</th>
<th>Definite</th>
<th>Ambiguous</th>
<th>No-Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>TWEET # 500</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DEFINITE</td>
<td>correct</td>
<td>243</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td>incorrect</td>
<td>33</td>
<td>17</td>
<td>10</td>
</tr>
<tr>
<td>AMBIGUOUS</td>
<td>correct</td>
<td>n/a</td>
<td>77</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td>incorrect</td>
<td>17</td>
<td>27</td>
<td>9</td>
</tr>
<tr>
<td>NO-LOCATION</td>
<td>correct</td>
<td>n/a</td>
<td>n/a</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>incorrect</td>
<td>7</td>
<td>13</td>
<td>n/a</td>
</tr>
</tbody>
</table>

Table 6.4: Ground Truth versus OzCT Geotagger Outcome (500 Tweets)
Evaluation

6.4 Evaluation

We evaluated the performance of the OzCT using two techniques. Our reasons for choosing these specific approaches are given in Section 3.3.4. The details of each evaluation technique is described next.

6.4.1 Evaluation against Human Judgement (Ground Truth)

The best method to evaluate the effectiveness of automated or semi-automated geotagging processes – such as identifying geo/non-geo references, specifying the geographic focus in the content and disambiguation of the results of the geo references – is to compare the results with manual human geotagging, which we use as the ground truth for evaluating the OzCT. Other geocoding research projects in various Web contexts benefit from manual human geocoding, which are described in Section 3.3.4.

To show this annotation step in practice, we provide examples of three tweets selected from the testing set below.

1. Posts valued at around $40,000, have been donated to farmers affected by the Yarrabin bushfire ⇒ Our human annotator detects a definite location: Yarrabin, NSW.

2. Residents of Cambridge Rd may be affected by #bushfire tonight ⇒ Our human annotator detects ambiguous locations: Cambridge Rd in various suburbs of VIC, TAS, SA, NSW and QLD states.

3. Cyclone unlikely to make landfall, as wet weather continues. ⇒ Our human annotator does not identify any location keyword and geotags this tweet as No-Location.

The output status of geotagging a geographical reference in a tweet content can be one of the following three: Definite, Ambiguous and No-Location, which might be detected as different geolocation references compared with the human judgement or detected them as other categories. For example, a definite location is detected as no-location or an ambiguous location is identified as INCORRECT. For no-location category, CORRECT means that OzCT did not detect any location the same as the ground truth; whereas, INCORRECT
applies to those non-geo entities that are mistakenly detected as definite or ambiguous location(s) by our system in which they were not geolocation by human judgement.

**Evaluation Measures**

Since toponym recognition is addressed as a natural language and text processing problem, two widely-used measures precision and recall are often suggested to evaluate such systems. In the context of tweet toponym recognition and resolution, we define recall \( R \) and precision \( P \) for ground truth identified locations \( GL \) and system-generated detected locations \( SL \) as:

\[
R(GL, SL) = \frac{|GL \cap SL|}{|GL|} \quad (6.1)
\]

\[
P(GL, SL) = \frac{|GL \cap SL|}{|SL|} \quad (6.2)
\]

In Equation 6.1, recall measures the proportion of identified locations in the ground truth that are correctly detected by the geotagger system. Whereas, in Equation 6.2, precision measures the proportion of all detected locations by geotagger system that are correct.

The results are shown in Table 6.4 and the metrics are compared in Table 6.7.

**6.4.2 Comparison with Existing GeoCoding Platforms**

To study the geographical focus of the OzCT we tested the same dataset with three geocoding systems and compared their results with the output of the OzCT for definite locations only. First, we used Stanford Named Entity Recognition (NER)\(^3\) as a baseline that identifies locations as well as other named entities such as organisations, persons, time and money. The NER systems identify entities in unseen text based on statistical models learnt from annotated text and do not require a geographical database (for example a gazetteer) to identify locations. Therefore, the comparison results are not biased towards particular locations list but are more generalised. However, different NER packages rely on different statistical algorithms.

\(^3\)http://nlp.stanford.edu/software/CRF-NER.shtml
to identify named entities and are useful in analysing static corpora rather than dynamic
text such as Web news and Twitter messages. As a result, different packages achieve various
and comparable results while parsing the same dataset. To overcome this issue, we chose
another widely used NLP software, the AlchemyAPI by the Orchestr8 company\(^4\) to have a
different set of named entity recognition results compared with the Stanford NER outcomes.

“AlchemyAPI is a text mining platform providing the most comprehensive set of semantic
analysis capabilities in the natural language processing field” ... “to read and understand
Web pages, text documents, emails, tweets, and other forms of content”. “AlchemyAPI
provides a scalable platform for analyzing Web pages, documents and tweets along with
flexible APIs for easy integration.” Finally, the Yahoo! GeoPlanet service\(^5\) is also chosen
to be the third evaluation approach against the OzCT due to its universal geographical
coverage. This platform aims to capture about six million geo-permanent named places
globally in a variety of languages by providing common vocabularies and grammar to facilitate
spatial interoperability and geographic discovery. Its “coverage varies from one country” to
another “including several hundred thousand unique administrative areas with half a million
variant names; several thousand historical administrative areas; over two million unique
settlements and suburbs, and millions of unique postal codes covering about 150 countries,
plus a significant number of Points of Interest, Colloquial Regions, Airports, Area Codes,
Time Zones, and Islands.”

**Evaluation Measures**

We also measure the **accuracy** of the aforementioned systems for the second type of evaluation
against other geocoding systems. A **Confusion Matrix** [Stehman, 1997] is used – a specific
table layout for visualising various aspects of an algorithm – to compare the outcome of
these systems for the four levels of geographical focus: *state*, *city*, *suburb* and *street*. In
brief, we compare how precisely the different systems can detect the geographical focus of
the location references. This comparison, however, is not applicable on the baseline (the
Stanford NER) as this package tags the tweets only by location name entity. We define

\(^4\)http://www.alchemyapi.com/

\(^5\)http://developer.yahoo.com/geo/geoplanet/
accuracy ($A$) in Equation 6.3 for correctly detected locations $SL$ by geotagger with a total of existing locations $T$ as follows:

$$A = \frac{SL}{T}$$

(6.3)

The results and the comparison graph are shown as a confusion matrix in Table 6.5 and Figure 6.2 respectively.

### 6.5 Results and Discussion

In this section, first, we discuss the results from two sets of experiments. Next, we propose an ontological toponym resolution for the OzCT geotagger’s ontological model that can be used as further step in resolving the ambiguity of location(s) results in the tweet content. Finally, several limitations of our approach are briefly discussed.

#### 6.5.1 Evaluation Results

We first evaluated our geotagger against the ground truth, which is a dataset of 500 tweets randomly sampled and manually geotagged. As displayed in Table 6.7, OzCT geotagger has achieved 81% recall for the detection of definite locations and its precision reaches 80.19% which results in the $F_1$-score of 80.40%. These results indicate that when there is a definite location reference in the tweet content, the geotagger detects them in over 80% of situations with the assurance of 80% correctness. In addition, our approach promises precision of 70.14% on no-location detection when there is no geographical reference in the tweet content; whereas, this number decreases to 57.46% for resolving ambiguous locations, which is the most complex status in geotagging tweets.

We studied the OzCT geotagger’s behaviour in more detail and identified two reasons for this considerable drop. One problem is with the gazetteers used. Google Map API is designed to retrieve the most relevant geographical location for the location entities and neglects the less frequent results. Another issue is that the toponym resolution phase could be further improved by using more heuristic rules. In addition, the proposed ontological model may be
Table 6.5: Accuracy of the Geographical Focus of Definite Locations in Three Geotaggers (Confusion Matrix)

<table>
<thead>
<tr>
<th>TWEET (#)</th>
<th>STATE</th>
<th>CITY</th>
<th>SUBURB</th>
<th>STREET</th>
</tr>
</thead>
<tbody>
<tr>
<td>300</td>
<td>110</td>
<td>85</td>
<td>64</td>
<td>41</td>
</tr>
<tr>
<td>AlchemyAPI NER</td>
<td>TRUE DETECTION (#)</td>
<td>36</td>
<td>28</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>ACCURACY (%)</td>
<td>32.72</td>
<td>30.58</td>
<td>10.93</td>
</tr>
<tr>
<td>Yahoo! GeoPlanet</td>
<td>TRUE DETECTION (#)</td>
<td>53</td>
<td>34</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>ACCURACY (%)</td>
<td>48.18</td>
<td>40.00</td>
<td>21.87</td>
</tr>
<tr>
<td>OzCT GEOTAGGER</td>
<td>TRUE DETECTION (#)</td>
<td>69</td>
<td>60</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>ACCURACY (%)</td>
<td>62.72</td>
<td>70.58</td>
<td>60.93</td>
</tr>
</tbody>
</table>

Table 6.6: Geotagging Time of Three geotagging Systems

<table>
<thead>
<tr>
<th>TWEET (#)</th>
<th>geotagging Time (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Per Tweet</td>
</tr>
<tr>
<td></td>
<td>AlchemyAPI NER</td>
</tr>
<tr>
<td>1</td>
<td>0.83</td>
</tr>
<tr>
<td>10</td>
<td>0.83</td>
</tr>
<tr>
<td>100</td>
<td>0.83</td>
</tr>
<tr>
<td>500</td>
<td>0.83</td>
</tr>
</tbody>
</table>
able to resolve such ambiguities.

In a separate experiment, for measuring the accuracy of the OzCT geotagger’s geographical focus, we tested our approach in recognising definite location with two different automated geocoding systems, Yahoo! GeoPlanet and AlchemyAPI NER were used to geotag the same dataset. As Figure 6.2 illustrates, all three platforms perform reasonably well to detect geographical references at the level of “state” and “city”, although their degree of coverage varies. Table 6.5 indicates that the AlchemyAPI NER detects 30% of tweets containing state and city references. Similarly, Yahoo! GeoPlanet service shows a better performance by an increase to 45% accuracy. These numbers significantly improve in our geotagger to 62.92%
Results and Discussion

Figure 6.3: Precision and Accuracy of Geographical Focus of OzCT Geotagger

and 70.58% for detection of states and cities respectively. From another perspective, AlchemyAPI and Yahoo! GeoPlanet perform poorly when there are suburb references in the tweet content (on average 15%), whereas, the OzCT geotagger detects suburbs in more than 60% of situations as well as it performs correctly in over half of the conditions where there are street-type references.

Time Analysis

From another perspective, the comparison of geotagging time for three systems is shown in Table 6.6, which explains some of the above behaviour of each system. We summarise our findings as follows.

1. The total time of geotagging tweets in OzCT geotagger is nearly two times slower than the two other geocoding systems.

2. The time of geotagging a tweet does not change for any number of tweets for AlchemyAPI NER and Yahoo! GeoPlanet platforms, whereas, this number decreases in OzCT geotagger by an increase of the tweet number. As a result, the total geotagging time of tweets is gradually decreased in our method.

The results can be interpreted in a way that in designing OzCT geotagger, we emphasised more on street and suburb detection by applying heuristic rules at the toponym resolution phase; whereas, in most geocoding systems, an identified location entity is considered as a
success and those systems do not query more detailed results. This could be because such systems are designed to detect the most detectable location in unstructured data and do not necessarily seek more details once a geolocation match is identified. Whereas, OzCT geotagger aims at detecting location references in the tweet content by the highest granularity. These views then explain why our method is almost two times slower than those systems due to time spend on more calculations.

Another explanation for such different behaviour could be that Yahoo! GeoPlanet and AlchemyAPI are not exclusively designed to specify location(s) in the tweet content considering the limitations with tweet (For example short length, noisy content). This can be the reason why the geotagging time of a tweet does not change for any number of tweets in these systems – AlchemyAPI NER (0.83 sec) and Yahoo! GeoPlanet (0.91 sec) – because these systems most likely perform the same process of location detection for all tweets. In comparison with OzCT geotagger, the geotagging time of a tweet decreases over time because our method performs location detection process in a more intelligent way. For example, OzCT geotagger stores the detected location keywords with the results from gazetteers’ search in a map during the run time. By finding the same location keyword(s) in next tweets, the program first searches the map before checking the gazetteers which results in reducing the overall time of geotagging process. This is also shown on Algorithm 4.1 in the second IF-ELSE statement.

We estimate that by evaluating OzCT geotagger against specific geocoding systems for tweets – to the best of our knowledge, there is no such tweet geotagging system so far – the results would be different.

Finally, we compared the precision of the OzCT geotagger for detecting three types of location with the accuracy of the geographical focus particularly for detecting a definite location given in Figure 6.3. The comparison indicates that almost two thirds of the testing dataset is geotagged as a definite location. Additionally, the OzCT geotagger is capable of location detection at different granularities (For example state, city, suburb and street) for almost a quarter of the datasets with definite locations. These figures are based on the 500 sample tweets and the results might be different for another dataset with larger or smaller scale. However, due to the nature of the testing dataset used in this research (a sample of
500 tweets out of over 22,000), we expect the overall behaviour of the OzCT geotagger will be similar for other tweet collections.

6.5.2 Ontological Toponym Resolution

As a further step in toponym resolution, we also develop an ontology model to apply semantic annotation of the tweet components, which is compared with relevant works in semantic annotation of social media in Table 6.3. Some ontologies are specifically created to model different kinds of social media such as user profile, social tagging, linking and other common user behaviour, which are described in Section 2.10.2.

In addition to the concepts and properties from existing data resources, we have also defined separate components for describing the particular properties for the ontology of the OzCrisisTracker application (ozon), which do not have corresponding existing concepts. For example, ozon:storyID and ozon:tweetCluster represent the classification in the training dataset of the Twitter messages. At the deeper analysis, we represent specific geographical location entities with the ozon namespace such as ozon:status, ozon:state, ozon:city, ozon:suburb.

Furthermore, a suitable formal language is required for the semantic annotation task. A discussion of this topic is provided in Section 2.3.3. We present the OzCT geotagger ontology in OWL/RDF language. Below, we briefly discuss how the proposed ontology model of tweets and applying SPARQL queries can be used to resolve the ambiguity of location results and improve the accuracy of the OzCT geotagger.

Take a Tweet from Table 6.2 as an example, a snapshot of its OWL representation using the ontology model is shown in Table 6.8. The namespaces used are shown in blue colour and string datatype are in red colour. In line 17, sioc:content indicates the content of the Tweet. In lines 19, 27 and 35, wgs84-pos:location specifies the points detected as the Tweets’ ambiguous locations (also shown in column four of Table 6.2).
Table 6.8: A Snapshot of OWL representation of a Tweet using Ontology Model
Results and Discussion

**Code 6.1: SPARQL Query**

```sql
1 SELECT ?subject ?object
2 WHERE { ozon:tweet1 wgs84_pos:location ?object }
```

**Code 6.2: SPARQL Query**

```sql
1 SELECT ?subject ?object ?state
2 WHERE { ozon:tweet1 wgs84_pos:location ?object .
3 ?object ozon:state ?state }
```

Complexity of SPARQL Queries

We also investigate a simple analysis on the complexity of the SPARQL queries. Given the Code 6.2, we can obtain the subject of 6 geographical points. Code 6.2 is performed once because only one location keyword **Bluff Road** is found with 6 answers. Next, Code 6.3 is performed 6 times as it queries the state (For example `ozon:state`) for each answer from the previous query. There are 5 answers for this query because there are 5 different states in the answer collection: **VIC, NSW, TAS, QLD and SA**. (Code 6.1, Line numbers 20, 28 and 36, not all answers are shown.) We then retrieve the street of each state. Code 6.4 is an example of querying the name of the streets (For example `ozon:street`) located in the **VIC** state, which is run 2 times because there are two answers for this query: **Bluff Road** and **St. Leonards**. Finally, a lookup of the street locations is applied for checking the distance of the streets found. Figure 6.4 is captured from Google Map showing the distance between these two streets. The overall number of SPARQL queries required to run is:

**Code 6.3: SPARQL Query**

```sql
1 SELECT ?subject ?object ?street
2 WHERE { ozon:tweet1 wgs84_pos:location ?object .
3   ?object ozon:state "VIC" .
4   ?object ozon:street ?street }
```
Results and Discussion

Figure 6.4: Google Map Representation of the Distance between Two Streets as an Example of Ontological Toponym Resolution

\[ O(n) = n + (n \times \#\text{location}) + (n \times \#\text{location} \times \#\text{state}) + (n \times \#\text{location} \times \#\text{state} \times \#\text{street}) \]

Looking at this simple example, we can conclude that whenever there is a state match between the results, a query for street level might resolve the ambiguity. As a result, the latter statement can be a new rule for semantic analysis of resolving the heterogeneity issues of ambiguous georeferences. Similar ontological analysis can apply on the output of the OzCT geotagger to discover more SPARQL queries and more semantic rules. However, in this work, we do not evaluate this step and ontological toponym resolution is an area for future research.
Summary

6.6 Summary

The overall objective of the present research is developing knowledge bases utilising ontologies for systematically representing data of two related areas of sustainability indicator sets and disaster management. In this chapter, we addressed the third focus of the research of developing an approach to identify and disambiguate the geographical references of emergency tweets which contain semantic heterogeneity issue in their georeferences. The OzCT geotagger automatically detects the location(s) mentioned in the content of tweets applying toponym resolution and toponym recognition methods. The output contains one of the three possibilities: definite, ambiguous and no-location. The present method also semantically annotates the tweet components utilising existing and new ontologies. We evaluated the method by comparing its output against human judgement as ground truth for a selection of 500 tweets that is manually geotagged by one of the authors prior to conducting the experiments. In addition, to study the geographical focus of the geotagger, we tested it using this dataset with other geotagging platforms including: The Alchemy API Named Entity Recognition and the Yahoo! Geo-Planet service.

The conclusions driven from result analysis were: (i) $F_1$ score of our system for detection of the definite locations was over 80%. It also showed on average 70% and 58% for no-location and ambiguous locations respectively. (ii) We also demonstrated that the accuracy of the geographical focus of the OzCT geotagger in detecting definite location, that was considerably higher than existing geotagging systems. While other platforms had lower coverage for suburb and street focus, our geotagger detected suburbs in more than 60% of situations and it performed correctly in nearly half of the conditions where there were street-type references. (iii) In designing the OzCT geotagger we emphasised more on street and suburb detection by applying heuristic rules at the toponym resolution phase; whereas, in most geotagging systems, an identified location entity is considered as a success and those systems do not query more detailed results. (iv) These views then explain why our method is almost two times slower than those systems due to time spent on more calculations.
Chapter 7

Conclusions and Future Work

"Two there are who are never satisfied – the lover of the world and the lover of knowledge."
—Rumi

The novelty of the present research is to explore semantic technologies to resolve heterogeneity issues of data representation for two related domains of sustainability and disaster management. Through addressing the research questions, we have focused on three specific aspects: (i) Constructing knowledge bases for resolving the issues of semantic heterogeneity of sustainability indicator sets, (ii) Developing a user interface for a reporting framework with two browsing mechanisms for heterogeneous sustainability indicators, and (iii) Constructing and evaluating a geotagger as a part of a decision making application in the context of disaster management, for detecting and disambiguating heterogeneous georeferences of the tweets’ content.

7.1 Solutions

With regards to the first research question, in the sustainability indicator sets domain, the semantic heterogeneity issue is identified through the representation of indicators and indicator sets. While different reporting organisations have various representations of indicators and their related concepts, presenting a uniform semantic model is a challenging task. During the
first phase of the research, we addressed this issue by putting forward two ontology designs for modelling entities and relationships in the sustainability indicators domain. The two design candidates, generic (GOSIS) and specific SOSIS, differ largely in terms of abstraction. We conclude that the specific design is preferable where the domain requirements demand a high degree of fidelity to seen frames of reference, while the GOSIS design offers greater reuse in contexts where unseen sets of indicators need to be added to the ontology in an ad-hoc fashion.

We took the output of the first research question and used it as the input for the second research question by applying the GOSIS design on CoS software to help end-user to browse heterogeneous indicator sets. The main challenge here is representing indicator sets, which are problematic from various reporting and navigation-related aspects. Two mechanisms for browsing indicator sets – Circular and Tabular views – were implemented for the development of the CoS user interface. We then conducted a user study of 26 participants. The questionnaire included 25 open and closed questions, which measured participants’ satisfaction of using the CoS software and its user-interface from two usability measures of perceived usefulness and perceived ease of use.

Analysing quantitative and qualitative results indicates several key findings. First, most end-users were satisfied with the CoS user interface, which simplified complex functionalities of reporting sustainability indicators. Second, the CoS software assisted users in understanding several features including: accomplishing tasks more quickly, increasing users’ productivity, and enhancing the effectiveness of sustainability assessment with the higher level of satisfaction. Next, the Circular view was the preferred interface to browse heterogeneous indicators by most participants. The Tabular view, however, had advantages of simplicity and easy search. Finally, the question of which view better assists users to use the CoS software, did not have a definite answer. It depends on which features of the CoS software users were choosing to exercise. Although, the Circular view is preferable for most reporting and navigation-related procedures, the Tabular view is also assumed to be the preferred alternative for basic search features. Our findings overall indicated that most end-users were satisfied with the CoS user interface, which simplified complex features of reporting with
sustainability indicators. Particularly, the Circular view assisted end-users in browsing heterogeneity of sustainability data by a better graphical display, ease of navigation and visual communication.

Addressing the third question required developing an approach to identify and disambiguate the heterogeneous geographical references within tweets in the context of disaster management, which is a related domain to sustainability. We constructed the OzCT geotagger for the Australian Crisis Tracker application, that automatically detects the location(s) mentioned in the tweets’ content by applying toponym resolution and toponym recognition methods. The output of OzCT geotagger was one of the three possibilities: definite, ambiguous and no-location. The present method also semantically annotated the tweet components utilising existing and new ontologies that was proposed to resolve the heterogeneity of georeferences in tweets’ content. Finally, OzCT was evaluated against human judgements (the ground truth) and two other geocoding systems. The results indicated that the accuracy of OzCT was significantly higher than other geocoding systems. The reason could be such systems are designed to detect the most available location in unstructured data and do not necessarily seek more details once a geolocation match was identified, whereas OzCT geotagger aimed at detecting location references at the finest granularity. In addition, we also proposed a semantic model that can improve disambiguation of georeferences by the use of existing and new ontologies.

7.2 Limitations

There are several limitations with the two ontology design candidates that we have developed. The contrasting results map intuitively to different requirements that can be said to underwrite the construction of the two ontology design candidates. Where an ontology needs to be precise and transparent – to faithfully represent, at a conceptual level, the specific conceptualisation of given indicator system – the approach adopted by SOSIS design is preferred, since it results in better $F$-measures where that system is used as a frame of reference. In contrast, GOSIS design better supports cases where the requirements emphasise ontology reuse with minimal cost of extension or refactoring, since it performs better against unseen
frames of reference. We also acknowledge there are cases where a compromise between these requirements might result in a hybrid of SOSIS and GOSIS designs. Indeed, such an option (maximising $F$-measures against both seen and unseen frames of reference) might be preferred where cost and time constraints permit.

Furthermore, we have identified several limitations with the OzCT. First, the use of the NER and PoS tagger slows the geotagging process. The total time of geotagging tweets in OzCT is nearly two times slower than the two other geocoding systems. This constraint becomes more critical when the system is deployed in a real time scenario and, for example, one hundred thousand tweets are in a queue and need to be geotagged with minimum delay. One solution is to investigate how effective is recognising the name entities before the toponym recognition phase. It may be possible to replace this step with some other techniques to accelerate the geotagging process. Second, the Google Map API is limited with the 20,000 queries per day, which causes the OzCT geotagger going to sleep for 24 hours. This constraint also generates serious complexities at the run time of geotagging procedure. Our solution to this was building a static gazetteer by storing the search results in a map (for example, an SQL table in the database) and using it for the similar queries. Another solution is using a new gazetteer with no query limitation, such as the Nominatim API\(^1\), and that is the future direction of this research. While the previous limitations are scalability issues, another challenge exists with the heuristic rules, which are detected from the training set. These patterns can be different when the dataset changes, for example, when the geotagger is used for identifying locations from Facebook messages. Therefore, static heuristic rules need to be changed and a more dynamic approach is required to detect the specific patterns from different resources, that suggests deeper understanding of the structure of other datasets.

### 7.3 Future Work

To recap, while reviewing the literature shows no coherent approach has been developed to tackle rival sustainability indicator systems in a “controlled” or “structured” setting, in this thesis we developed well-design ontologies that can be effectively used in reporting practices.

\(^1\)http://www.nominatim.org/
for browsing and navigation purposes. Similarly, studies of social media in the context of disaster management do not adequately address the disambiguation of georeference of tweets’ content in an “uncontrolled” or “unstructured” setting, so we developed a geotagger with the capability of higher coverage and significant accuracy at detecting absolute locations with the use of geographical knowledge bases and semantic modelling. We anticipate future work on each question along several lines and perhaps semantic heterogeneity can never be completely solved, but various technical approaches can alleviate some of the issues for users.

With regards to the first research question, first, we think that ROMEO can be linked to METHONTOLOGY in a more systematic way, to guide ontology development from requirements through to evaluation and selection. Second, ROMEO itself can be extended through the sorts of quantitative procedures we apply here. Finally, further work can be undertaken to incorporate additional sustainability indicators systems, and to further refine the candidate OSIS ontologies presented in this research. We also anticipate the possibility of blending both approaches in future. More generally, we show that both METHONTOLOGY and ROMEO can be productively used to guide the design and evaluation of domain-level ontologies, and that quantitative measures such as the $F$-measure can be used to develop heuristics for preferring one ontology candidate to another, given a set of requirements and frames of reference.

With regards to the second research question, we think a combined user interface of visual and textual methods for browsing sustainability indicator sets would greatly improve the data representation issues. In addition, a sustainability reporting framework requires a concrete guidance or a tour to train less-expert users and simplify the complex concepts of sustainability domain for them. We then suggest the use of an interactive interface that associates with the appropriate guidelines to simplify the complexities of working with such a sustainability reporting interface.

With regards to the third research question, the accuracy of the OzCT can be improved using more accurate gazetteers (such as Nominatim API). Furthermore, the OzCT geotagger precision may be enhanced by applying external factors, such as user location, retweeted frequency and considering the time zone of the tweets. IBM Research – Australia plans
deployment of the OzCrisisTracker 2.0 for the Australian Bushfire and Flood season 2015. It is expected that in the real-time scenarios, the output of the OzCT module will enhance the whole system’s performance by (i) Predicting different disaster events based on geolocation detected in tweets’ content, and (ii) Improving the clustering process of tweets by adding a geolocation feature, which will reduce the information overload for the clustering algorithm. However, considering the overall slowness of the geotagging time, we will apply some changes on the current system for the real-time situations. For example, the timely name entity recognition process will be switched off and the speed of location detection will be improved greatly once Google Map API is replaced with a static gazetteer. Furthermore, data resources such as Flickr, Facebook and Youtube as the feeder for OzCrisisTracker 2.0 will be also taken into account and the OzCT geotagger is expected to geotag such complex multimedia messages with some further development in analysing multimedia messages. Finally, we suggest the presented semantic annotation of tweets in this research utilising existing and new ontologies, can be used for conceptual analysis of tweets’ content and will resolve the ambiguity of tweet locations to significantly improve the geotagger performance. Moreover, our suggested ontological model for tweets will enhance the integration of the OzCrisisTracker system with other applications in the disaster management domain.

7.4 Final Remarks

While more effort is required to investigate technical solutions to problems of semantic heterogeneity for the domains of sustainability and disaster management, this thesis has addressed some of the data modelling and representation issues of heterogeneous resources of both structured (sustainability indicator sets) and semi or unstructured (social media) through a series of semantic technologies: ontology engineering, ontology design patterns, user interface models and geotagging tweets. In resolving heterogeneity issues within these domains, we emphasise the importance of flexibility of a generic ontology design (GOSIS), usability of a browsing interface (Circular view) and accuracy of a geotagger for tweets (OzCT).
Appendix A

OWL Code of OSIS Designs

Here, we present OWL code of two OSIS design candidates: GOSIS and SOSIS in Table A.1 to A.6.
<?xml version="1.0"?>
<!DOCTYPE Ontology [ 
<!ENTITY xsd "http://www.w3.org/2001/XMLSchema#" >
<!ENTITY xml "http://www.w3.org/XML/1998/namespace" >
<!ENTITY rdfs "http://www.w3.org/2000/01/rdf-schema#" >
<!ENTITY rdf "http://www.w3.org/1999/02/22-rdf-syntax-ns#" >
]>
<Ontology xmlns="http://www.w3.org/2002/07/owl#"
xml:base="http://www.semanticweb.org/ontologies/2012/08/GOSIS"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
ontologyIRI="http://www.semanticweb.org/ontologies/2012/08/GOSIS">
<Prefix name="" IRI="http://www.w3.org/2002/07/owl#"/>
<Prefix name="dc" IRI="http://purl.org/dc/elements/1.1/"/>
<Prefix name="owl" IRI="http://www.w3.org/2002/07/owl#"/>
<Prefix name="rdf" IRI="http://www.w3.org/1999/02/22-rdf-syntax-ns#"/>
<Prefix name="xsd" IRI="http://www.w3.org/2001/XMLSchema#"/>
<Prefix name="osis" IRI="http://www.cs.rmit.edu.au/knowledgebase/ontology/OSIS/"/>
<Prefix name="rdfs" IRI="http://www.w3.org/2000/01/rdf-schema#"/>
<Import>http://purl.org/dc/terms/</Import>
<Import>http://purl.org/dc/elements/1.1/</Import>
<Import>http://purl.org/dc/dcmitype/</Import>
<Declaration>
<Class IRI="#AAA"/>
</Declaration>
<Declaration>
<Class IRI="#Category"/>
</Declaration>
<Declaration>
<Class IRI="#Description"/>
</Declaration>
<Declaration>
<Class IRI="#Indicator"/>
</Declaration>
<Declaration>
<Class IRI="#IndicatorSet"/>
</Declaration>
<Declaration>
<Class IRI="#Issue"/>
</Declaration>
<Declaration>
<Class IRI="#Reference"/>
</Declaration>
</Ontology>
Table A.2: GOSIS Design Ontology OWL Code - Part 2
Table A.3: GOSIS Design Ontology OWL Code - Part 3
<table>
<thead>
<tr>
<th>Line</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>&lt;?xml version=&quot;1.0&quot;?&gt;</td>
</tr>
<tr>
<td>2</td>
<td>&lt;!DOCTYPE Ontology [</td>
</tr>
<tr>
<td>3</td>
<td>&lt;!ENTITY xsd &quot;http://www.w3.org/2001/XMLSchema#&quot; &gt;</td>
</tr>
<tr>
<td>4</td>
<td>&lt;!ENTITY xml &quot;http://www.w3.org/XML/1998/namespace&quot; &gt;</td>
</tr>
<tr>
<td>5</td>
<td>&lt;!ENTITY rdfs &quot;http://www.w3.org/2000/01/rdf-schema#&quot; &gt;</td>
</tr>
<tr>
<td>6</td>
<td>&lt;!ENTITY rdf &quot;http://www.w3.org/1999/02/22-rdf-syntax-ns#&quot; &gt;</td>
</tr>
<tr>
<td>7</td>
<td>]&gt;</td>
</tr>
<tr>
<td>8</td>
<td>&lt;Ontology xmlns=&quot;<a href="http://www.w3.org/2002/07/owl#">http://www.w3.org/2002/07/owl#</a>&quot;</td>
</tr>
<tr>
<td>9</td>
<td>xml:base=&quot;<a href="http://www.semanticweb.org/ontologies/2012/8/SOSIS.owl">http://www.semanticweb.org/ontologies/2012/8/SOSIS.owl</a>&quot;</td>
</tr>
<tr>
<td>10</td>
<td>xmlns:rdfs=&quot;<a href="http://www.w3.org/2000/01/rdf-schema#">http://www.w3.org/2000/01/rdf-schema#</a>&quot;</td>
</tr>
<tr>
<td>11</td>
<td>xmlns:xsd=&quot;<a href="http://www.w3.org/2001/XMLSchema#">http://www.w3.org/2001/XMLSchema#</a>&quot;</td>
</tr>
<tr>
<td>12</td>
<td>xmlns:rdf=&quot;<a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#">http://www.w3.org/1999/02/22-rdf-syntax-ns#</a>&quot;</td>
</tr>
<tr>
<td>14</td>
<td>ontologyIRI=&quot;<a href="http://www.semanticweb.org/ontologies/2012/8/SOSIS.owl%22%3E">http://www.semanticweb.org/ontologies/2012/8/SOSIS.owl&quot;&gt;</a></td>
</tr>
<tr>
<td>15</td>
<td>&lt;Prefix name=&quot;xsd&quot; IRI=&quot;http://www.w3.org/2001/XMLSchema#&quot;/&gt;</td>
</tr>
<tr>
<td>16</td>
<td>&lt;Prefix name=&quot;owl&quot; IRI=&quot;http://www.w3.org/2002/07/owl#&quot;/&gt;</td>
</tr>
<tr>
<td>17</td>
<td>&lt;Prefix name=&quot;gri&quot; IRI=&quot;https://www.globalreporting.org&quot;/&gt;</td>
</tr>
<tr>
<td>18</td>
<td>&lt;Prefix name=&quot;osis&quot;</td>
</tr>
<tr>
<td>20</td>
<td>&lt;Prefix name=&quot;oecd&quot; IRI=&quot;http://www.oecd.org&quot;/&gt;</td>
</tr>
<tr>
<td>21</td>
<td>&lt;Prefix name=&quot;dc&quot; IRI=&quot;http://purl.org/dc/elements/1.1&quot;/&gt;</td>
</tr>
<tr>
<td>22</td>
<td>&lt;Import&gt;<a href="http://purl.org/dc/dcmitype/">http://purl.org/dc/dcmitype/</a>&lt;/Import&gt;</td>
</tr>
<tr>
<td>23</td>
<td>&lt;Import&gt;<a href="http://purl.org/dc/elements/1.1/">http://purl.org/dc/elements/1.1/</a>&lt;/Import&gt;</td>
</tr>
<tr>
<td>24</td>
<td>&lt;Import&gt;<a href="http://purl.org/dc/terms/">http://purl.org/dc/terms/</a>&lt;/Import&gt;</td>
</tr>
<tr>
<td>25</td>
<td>&lt;Import&gt;<a href="http://purl.org/dc/dcmitype/">http://purl.org/dc/dcmitype/</a>&lt;/Import&gt;</td>
</tr>
<tr>
<td>26</td>
<td>&lt;/Declaration&gt;</td>
</tr>
<tr>
<td>27</td>
<td>&lt;Declaration&gt;</td>
</tr>
<tr>
<td>28</td>
<td>&lt;Class IRI=&quot;#Category&quot;/&gt;</td>
</tr>
<tr>
<td>29</td>
<td>&lt;/Declaration&gt;</td>
</tr>
<tr>
<td>30</td>
<td>&lt;Declaration&gt;</td>
</tr>
<tr>
<td>31</td>
<td>&lt;Class IRI=&quot;#Description&quot;/&gt;</td>
</tr>
<tr>
<td>32</td>
<td>&lt;/Declaration&gt;</td>
</tr>
<tr>
<td>33</td>
<td>&lt;Declaration&gt;</td>
</tr>
<tr>
<td>34</td>
<td>&lt;Class IRI=&quot;#GRI_Aspect&quot;/&gt;</td>
</tr>
<tr>
<td>35</td>
<td>&lt;/Declaration&gt;</td>
</tr>
<tr>
<td>36</td>
<td>&lt;Declaration&gt;</td>
</tr>
<tr>
<td>37</td>
<td>&lt;Class IRI=&quot;#GRI_Description&quot;/&gt;</td>
</tr>
<tr>
<td>38</td>
<td>&lt;/Declaration&gt;</td>
</tr>
<tr>
<td>39</td>
<td>&lt;Declaration&gt;</td>
</tr>
<tr>
<td>40</td>
<td>&lt;Class IRI=&quot;#GRI_Indicator&quot;/&gt;</td>
</tr>
<tr>
<td>41</td>
<td>&lt;/Declaration&gt;</td>
</tr>
<tr>
<td>42</td>
<td>&lt;Declaration&gt;</td>
</tr>
<tr>
<td>43</td>
<td>&lt;Class IRI=&quot;#Indicator&quot;/&gt;</td>
</tr>
<tr>
<td>44</td>
<td>&lt;/Declaration&gt;</td>
</tr>
<tr>
<td>45</td>
<td>&lt;Declaration&gt;</td>
</tr>
<tr>
<td>46</td>
<td>&lt;Class IRI=&quot;#OECD_Description&quot;/&gt;</td>
</tr>
<tr>
<td>47</td>
<td>&lt;/Declaration&gt;</td>
</tr>
<tr>
<td>48</td>
<td>&lt;Declaration&gt;</td>
</tr>
<tr>
<td>49</td>
<td>&lt;Class IRI=&quot;#OECD_Indicator&quot;/&gt;</td>
</tr>
<tr>
<td>50</td>
<td>&lt;/Declaration&gt;</td>
</tr>
<tr>
<td>51</td>
<td>&lt;Declaration&gt;</td>
</tr>
<tr>
<td>52</td>
<td>&lt;Class IRI=&quot;#OECD_Theme&quot;/&gt;</td>
</tr>
<tr>
<td>53</td>
<td>&lt;/Declaration&gt;</td>
</tr>
<tr>
<td>54</td>
<td>&lt;Declaration&gt;</td>
</tr>
<tr>
<td>55</td>
<td>&lt;Class IRI=&quot;#Reference&quot;/&gt;</td>
</tr>
<tr>
<td>56</td>
<td>&lt;/Declaration&gt;</td>
</tr>
</tbody>
</table>

Table A.4: SOSIS Design Ontology OWL Code - Part 1
<table>
<thead>
<tr>
<th>Line</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><code>&lt;Declaration&gt;</code></td>
</tr>
<tr>
<td>2</td>
<td><code>&lt;ObjectProperty abbreviatedIRI=&quot;dc:Title&quot;/&gt;</code></td>
</tr>
<tr>
<td>3</td>
<td><code>&lt;/Declaration&gt;</code></td>
</tr>
<tr>
<td>4</td>
<td><code>&lt;Declaration&gt;</code></td>
</tr>
<tr>
<td>5</td>
<td><code>&lt;ObjectProperty abbreviatedIRI=&quot;dc:Type&quot;/&gt;</code></td>
</tr>
<tr>
<td>6</td>
<td><code>&lt;/Declaration&gt;</code></td>
</tr>
<tr>
<td>7</td>
<td><code>&lt;Declaration&gt;</code></td>
</tr>
<tr>
<td>8</td>
<td><code>&lt;ObjectProperty abbreviatedIRI=&quot;osis:hasCategory&quot;/&gt;</code></td>
</tr>
<tr>
<td>9</td>
<td><code>&lt;/Declaration&gt;</code></td>
</tr>
<tr>
<td>10</td>
<td><code>&lt;Declaration&gt;</code></td>
</tr>
<tr>
<td>11</td>
<td><code>&lt;ObjectProperty abbreviatedIRI=&quot;osis:hasReference&quot;/&gt;</code></td>
</tr>
<tr>
<td>12</td>
<td><code>&lt;/Declaration&gt;</code></td>
</tr>
<tr>
<td>13</td>
<td><code>&lt;Declaration&gt;</code></td>
</tr>
<tr>
<td>14</td>
<td><code>&lt;ObjectProperty abbreviatedIRI=&quot;osis:hasUnitofMeasurement&quot;/&gt;</code></td>
</tr>
<tr>
<td>15</td>
<td><code>&lt;/Declaration&gt;</code></td>
</tr>
<tr>
<td>16</td>
<td><code>&lt;Declaration&gt;</code></td>
</tr>
<tr>
<td>17</td>
<td><code>&lt;ObjectProperty abbreviatedIRI=&quot;osis:instanceOf&quot;/&gt;</code></td>
</tr>
<tr>
<td>18</td>
<td><code>&lt;/Declaration&gt;</code></td>
</tr>
<tr>
<td>19</td>
<td><code>&lt;Declaration&gt;</code></td>
</tr>
<tr>
<td>20</td>
<td><code>&lt;ObjectProperty abbreviatedIRI=&quot;oecd:hasDefinition&quot;/&gt;</code></td>
</tr>
<tr>
<td>21</td>
<td><code>&lt;/Declaration&gt;</code></td>
</tr>
<tr>
<td>22</td>
<td><code>&lt;Declaration&gt;</code></td>
</tr>
<tr>
<td>23</td>
<td><code>&lt;ObjectProperty abbreviatedIRI=&quot;oecd:hasInformation&quot;/&gt;</code></td>
</tr>
<tr>
<td>24</td>
<td><code>&lt;/Declaration&gt;</code></td>
</tr>
<tr>
<td>25</td>
<td><code>&lt;Declaration&gt;</code></td>
</tr>
<tr>
<td>26</td>
<td><code>&lt;ObjectProperty abbreviatedIRI=&quot;oecd:hasTheme&quot;/&gt;</code></td>
</tr>
<tr>
<td>27</td>
<td><code>&lt;/Declaration&gt;</code></td>
</tr>
<tr>
<td>28</td>
<td><code>&lt;Declaration&gt;</code></td>
</tr>
<tr>
<td>29</td>
<td><code>&lt;ObjectProperty abbreviatedIRI=&quot;gri:hasAspect&quot;/&gt;</code></td>
</tr>
<tr>
<td>30</td>
<td><code>&lt;/Declaration&gt;</code></td>
</tr>
<tr>
<td>31</td>
<td><code>&lt;Declaration&gt;</code></td>
</tr>
<tr>
<td>32</td>
<td><code>&lt;ObjectProperty abbreviatedIRI=&quot;gri:hasCompilation&quot;/&gt;</code></td>
</tr>
<tr>
<td>33</td>
<td><code>&lt;/Declaration&gt;</code></td>
</tr>
<tr>
<td>34</td>
<td><code>&lt;Declaration&gt;</code></td>
</tr>
<tr>
<td>35</td>
<td><code>&lt;ObjectProperty abbreviatedIRI=&quot;gri:hasDefinition&quot;/&gt;</code></td>
</tr>
<tr>
<td>36</td>
<td><code>&lt;/Declaration&gt;</code></td>
</tr>
<tr>
<td>37</td>
<td><code>&lt;Declaration&gt;</code></td>
</tr>
<tr>
<td>38</td>
<td><code>&lt;ObjectProperty abbreviatedIRI=&quot;gri:hasDocumentation&quot;/&gt;</code></td>
</tr>
<tr>
<td>39</td>
<td><code>&lt;/Declaration&gt;</code></td>
</tr>
<tr>
<td>40</td>
<td><code>&lt;Declaration&gt;</code></td>
</tr>
<tr>
<td>41</td>
<td><code>&lt;ObjectProperty abbreviatedIRI=&quot;gri:hasRelevance&quot;/&gt;</code></td>
</tr>
<tr>
<td>42</td>
<td><code>&lt;/Declaration&gt;</code></td>
</tr>
</tbody>
</table>

Table A.5: SOSIS Design Ontology OWL Code - Part 2
Table A.6: SOSIS Design Ontology OWL Code - Part 3

```
<SubClassOf>
  <Class IRI="#GRI_Aspect"/>
  <Class IRI="#Category"/>
</SubClassOf>

<SubClassOf>
  <Class IRI="#GRI_Description"/>
  <Class IRI="#Description"/>
</SubClassOf>

<SubClassOf>
  <Class IRI="#GRI_Indicator"/>
  <Class IRI="#Indicator"/>
</SubClassOf>

<SubClassOf>
  <Class IRI="#OECD_Description"/>
  <Class IRI="#Description"/>
</SubClassOf>

<SubClassOf>
  <Class IRI="#OECD_Indicator"/>
  <Class IRI="#Indicator"/>
</SubClassOf>

<SubClassOf>
  <Class IRI="#OECD_Theme"/>
  <Class IRI="#Category"/>
</SubClassOf>

<ObjectPropertyDomain>
  <ObjectProperty abbreviatedIRI="oecd:hasDefinition"/>
  <Class IRI="#OECD_Indicator"/>
</ObjectPropertyDomain>

<ObjectPropertyDomain>
  <ObjectProperty abbreviatedIRI="oecd:hasInformation"/>
  <Class IRI="#OECD_Indicator"/>
</ObjectPropertyDomain>

<ObjectPropertyDomain>
  <ObjectProperty abbreviatedIRI="oecd:hasTheme"/>
  <Class IRI="#OECD_Indicator"/>
</ObjectPropertyDomain>

<ObjectPropertyDomain>
  <ObjectProperty abbreviatedIRI="gri:hasAspect"/>
  <Class IRI="#GRI_Indicator"/>
</ObjectPropertyDomain>

<ObjectPropertyDomain>
  <ObjectProperty abbreviatedIRI="gri:hasCompilation"/>
  <Class IRI="#GRI_Indicator"/>
</ObjectPropertyDomain>

<ObjectPropertyDomain>
  <ObjectProperty abbreviatedIRI="gri:hasDefinition"/>
  <Class IRI="#GRI_Indicator"/>
</ObjectPropertyDomain>

<ObjectPropertyDomain>
  <ObjectProperty abbreviatedIRI="gri:hasDocumentation"/>
  <Class IRI="#GRI_Indicator"/>
</ObjectPropertyDomain>

<ObjectPropertyDomain>
  <ObjectProperty abbreviatedIRI="gri:hasRelevance"/>
  <Class IRI="#GRI_Indicator"/>
</ObjectPropertyDomain>

</Ontology>
```
Appendix B

CoS User Study - Ethics

For conducting the user study to evaluate CoS software in Chapter 5, there was a need for ethics approval for research projects involving humans. Prior to conducting the user study experiment, we gained the approval from the college Human Ethics Advisory Network (CHEAN). We provided a plain language statement of the experiment, whose link was presented to the participants at the beginning of the questionnaire. This statement – shown in Figures B.1 to B.4 – describes the aim of the project, the people who are involved, the tasks needed to be completed, the rights of the participants and the contacts information of people in charge.
Figure B.1: Ethics Approval for CoS Research
Invitation to participate in a research project

Project information statement

Project Title:
- Accounting for Sustainability: Developing an Integrated Approach for Sustainability Assessments

Investigators:
- Prof. Paul James (Chief Investigator: Director, Global Cities Institute, RMIT University, paul.james@rmit.edu.au, 9925-2500)
- Dr. Andy Scerri (Research Fellow, Global Cities Institute, RMIT University, andy.scerri@rmit.edu.au, 9925-1946)
- Dr. Liam Magee (Research Fellow, Globalism Institute, RMIT University, liam.magee@rmit.edu.au, 9925-2637)
- Dr. Sarah Hickmott (Research Fellow, Globalism Institute, RMIT University, sarah.hickmott@gmail.com)
- Assoc. Prof. James Thom (Research Fellow, Globalism Institute, RMIT University, james.thom@rmit.edu.au, 9925-2992)
- Lida Ghaheemanlou (PhD candidate, Globalism Institute, RMIT University, lida.ghahremanlou@rmit.edu.au, 9925-2758)

Dear Participant,
You are invited to participate in a research project being conducted by RMIT University. This information sheet describes the project in straightforward language, or 'plain English'. Please read this sheet carefully and be confident that you understand its contents before deciding whether to participate. If you have any questions about the project, please ask one of the investigators.

Who is involved in this research project? Why is it being conducted?
This Australian Research Council project is being conducted by a team at RMIT University, in partnership with and partially funded by FujiXerox Australia, Cambridge International College, Microsoft Australia, Common Ground Publishing, Angusta Systems, the City of Melbourne, the City of Vancouver and the Australian Government. The project is supervised by Prof. Paul James, Prof. Lin Padgham, Assoc. Prof. James Thom and Assoc. Prof. Hepu Deng. Dr. Andy Scerri, Liam Magee and Sarah Hickmott, and PhD candidates Lida Ghaheemanlou and David The form the research team.

The project investigates how companies, local governments and communities assess and report on sustainability. It is also piloting an alternative approach to understanding critical issues around sustainability for particular communities. It has been approved by the RMIT Human Research Ethics Committee.

Why have you been approached?
We are seeking users to provide feedback on a questionnaire designed to measure the overall usefulness of the Circles of Sustainability software and the ease of use of two ways of browsing sustainability indicators provided in the interface. We are inviting participants who are familiar with sustainability methods and theories. You are likely to

Figure B.2: Plain Language Statement for the CoS User Study Evaluation - Page 1
have been invited through recommendation of people associated with the Global Compact Cities Program. Since the purpose of this research is to test the usability of the questionnaire, in some specific cases we may approach individuals based on the recommendations of other participants or researchers, using what is often termed the “snowball” approach to recruiting participants.

What is the project about? What are the questions being addressed?

The first focus in the PhD research engaged with this project is representing semantic heterogeneity of sustainability indicators. An ontology model is developed as the computational solution to this problem. The ontology model is then applied on the Circles of Sustainability software and a user-study is designed to address the second research question on what mechanisms can be used in a sustainability reporting framework to help end-users navigating and browsing semantic heterogeneity of the sustainability indicators.

Accordingly, two objectives are identified for the user-study and the questions from perspectives of two participants (expert and non-expert).

1. The first goal aims at evaluating the usefulness of the software overall.
2. The second goal looks into ease of use of two ways of browsing indicators provided in the interface: Table View vs. Circular View.

If I agree to participate, what will I be required to do?

You will be given a set of instruction to do an experiment. Two tasks are expected from you to accomplish the experiment.

1. Adding relevant indicators to issues using both views.
2. Assessing the sample project using the added indicators against Circles of Sustainability.

You will be also to complete a questionnaire, comprising 23 close questions, and 2 open ended text questions. You will be asked to indicate your level of satisfaction, on a scale of 1-5. We estimate completing the experiment and questionnaire will take approximately 20-30 minutes.

What are the risks or disadvantages associated with participation?

We don’t believe participation in this research poses any risks outside of normal day-to-day activities. The questionnaire does however aim to measure your personal sense of wellbeing as well as general attitudes towards local and global sustainability. If you are unduly concerned about your responses to any of the questionnaire items or if you find participation in the project distressing, you should contact Professor Paul James as soon as convenient. Paul James will discuss your concerns with you confidentially and suggest appropriate follow-up, if necessary.

What are the benefits associated with participation?

There are no direct benefits associated with participation in this research. You may find reflection upon the themes and prompts suggested by questionnaire interesting and provoking.
What will happen to the information I provide?

Information gathered through the questionnaire will be stored on secure computer systems managed by RMIT University. Your response to the questionnaire should not include any identifying information; if in response to open-ended questions you include obvious identifying information, this will be removed prior to data entry.

Results from the questionnaire may be published in academic journal articles or conference proceedings, books, reports to local government or in other appropriate forums. Such results will be aggregated, and will not contain any identifying information. This data will be kept securely at RMIT for a period of 5 years before being destroyed.

Any information that you provide can be disclosed only if (1) it is to protect you or others from harm, (2) a court order is produced, or (3) you provide the researchers with written permission.

Because of the nature of data collection, we are not obtaining written informed consent from you. Instead, we assume that you have given consent by your completion and return of the questionnaires.

What are my rights as a participant?

As a participant in this research, you may insist upon:

- The right to withdraw your participation at any time, without prejudice.
- The right to have any unprocessed data withdrawn and destroyed, provided it can be reliably identified, and provided that so doing does not increase the risk for the participant.
- The right to have any questions answered at any time.

Whom should I contact if I have any questions?

You may contact any of the researchers directly involved in this part of the project. Contact details are provided under the list of the investigators at the start of this document.

What other issues should I be aware of before deciding whether to participate?

There are no other issues we think you need to be aware of before deciding whether to participate in the research.

If you have any questions or concerns about this research please contact my supervisors:

A/Prof. James Thom (james.thom@rmit.edu.au)
Dr. Liam Magee (liam.magee@rmit.edu.au)

Yours sincerely,

Lida Ghahremanlou

Figure B.4: Plain Language Statement for the CoS User Study Evaluation - Page 3
Appendix C

CoS User Study - Questionnaire

The questionnaire of the user study conducted in Chapter 5, is divided into the three sections. The first contains 5 closed questions with a scale of 1-4 and assesses the background knowledge of the participants about sustainability concepts and reporting organisations. Participants are asked to answer this section before using the software and answering the other two sections. The second section of the questionnaire contains 6 questions in a 5-point Likert Scale and assesses the overall usability of the software. The third section contains two subsections for measuring “perceived ease of use” for Tabular and Circular views respectively. Figures C.1 to C.5 show the questions.

**NOTE:** During the user study experiment, the Tabular view was referred to as the Table view in the instructions given to users.
Appendix C. COS User Study - Questionnaire

Dear Paul,

Re: Human Research Ethics Application - Register Number CHEAN A 3006558-08/11

The Deputy Chair of the Design and Social Context College Human Ethics Advisory Network (CHEAN), Prof Joseph Sinexa, received your ethics application for the following research project:

Accounting for Sustainability: Developing an Integrated Approach for Sustainability Assessments

I am pleased to inform you that your application has been approved at a Low Risk classification. This approval will be ratified at the meeting on 18 September 2011 and will be reported to the University Human Research Ethics Committee for noting.

Your ethics approval expires on 31 December 2013.

Please note that all research data should be stored on University Network systems. These systems provide high levels of manageable security and data encryption, can provide secure remote access, are backed up on a regular basis and can provide Disaster Recovery processes should a large scale incident occur. The use of portable devices such as CDs and memory sticks is invalid for archiving, data transport where necessary and some works in progress. The authoritative copy of all research data should reside on appropriate network systems, and the Principal Investigator is responsible for the retention and storage of the original data pertaining to the project for a minimum period of five years.

You are reminded that an Annual Report is mandatory and should be forwarded to the College Ethics Sub-committee Secretary by mid-January 2012. This report is available at http://www.rmit.edu.au/humres/pdf/70gdg%eBDgdp or can be located by following the link under Policy at http://www.rmit.edu.au/dsc/hr.

Should you have any queries regarding your application please seek advice from the Deputy Chair of the College Human Ethics Advisory Network (CHEAN) Prof Joseph Sinexa on (03) 9925 1744, joseph.sinexa@rmit.edu.au or contact Lisa Mann on (03) 9925 2574, lisa.mann@rmit.edu.au

On behalf of the DSC College Human Ethics Advisory Network I wish you well in your research.

Yours sincerely,

Lisa Mann
Ethics Officer
DSC College Human Ethics Advisory Network (CHEAN)

Figure C.1: Consent Form
Welcome to the Circles of Sustainability survey. Please first read the background material and answer Questions 1 to 5.

In response to measuring and maintaining the sustainability status of a system, a number of indicator systems have been developed and are in use today. These include generalized reporting standards, such as those developed by the Global Reporting Initiative (GRI), the Organization for Economic Co-operation and Development (OECD) and the United Nations Statistics Division (UN Social Indicators).

**Q1. How familiar are you with sustainability concepts, theories, approaches and methods?**
- Very familiar
- Somewhat familiar
- A little familiar
- Not familiar

**Q2. How familiar are you with sustainability indicators of Global Reporting Initiative (GRI)?**
- Very familiar
- Somewhat familiar
- A little familiar
- Not familiar

**Q3. How familiar are you with sustainability indicators of Organisation of Economic Co-operation and Development (OECD)?**
- Very familiar
- Somewhat familiar
- A little familiar
- Not familiar

**Q4. How familiar are you with social indicators of United Nations Statistics Division (UN)?**
- Very familiar
- Somewhat familiar
- A little familiar
- Not familiar
Sustainability reporting framework are used to assess a project against standardized indicator sets. Circles of Sustainability (CoS) is a practical user-friendly tool that project administrators can utilize to implement and track their sustainable project with measurable outcomes. CoS methodology contains the concepts such as: creating project, defining critical issues, adding indicators and assessing the sample project against circles of sustainability.

<table>
<thead>
<tr>
<th>Q5. How familiar are you with the Circles of Sustainability concepts and its methodology?</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐ Very familiar</td>
</tr>
<tr>
<td>☐ Somewhat familiar</td>
</tr>
<tr>
<td>☐ A little familiar</td>
</tr>
<tr>
<td>☐ Not familiar</td>
</tr>
</tbody>
</table>

*Figure C.3: Section 1 - Background Knowledge of the Participants - Page 2*
Figure C.4: Section 2 - Overall Usefulness of the CoS Software

Table: Please answer questions from 6 to 11 related to overall usefulness of the CoS software.

<table>
<thead>
<tr>
<th>Q6. Using the CoS software in sustainability reporting would enable me to accomplish tasks more quickly.</th>
</tr>
</thead>
<tbody>
<tr>
<td>extremely likely</td>
</tr>
<tr>
<td>☐</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>extremely likely</td>
</tr>
<tr>
<td>☐</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Q8. Using the CoS software in sustainability assessment would increase my productivity.</th>
</tr>
</thead>
<tbody>
<tr>
<td>extremely likely</td>
</tr>
<tr>
<td>☐</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Q9. Using the CoS software would enhance my effectiveness in sustainability assessment.</th>
</tr>
</thead>
<tbody>
<tr>
<td>extremely likely</td>
</tr>
<tr>
<td>☐</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Q10. Using the CoS software would make it easier to do sustainability assessment.</th>
</tr>
</thead>
<tbody>
<tr>
<td>extremely likely</td>
</tr>
<tr>
<td>☐</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Q11. I would find the CoS software useful in sustainability assessment.</th>
</tr>
</thead>
<tbody>
<tr>
<td>extremely likely</td>
</tr>
<tr>
<td>☐</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Q12. Can you imagine using the software in everyday sustainability assessment and reporting practices? What could be done or added to make it more useful?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>


Figure C.5: Section 3 - Ease of Use of Browsing Mechanisms of Indicator Sets
Appendix D

CoS User Study - Scripts

In order to guide the participants in performing the tasks, we also provided a set of instructions for using the CoS software and accomplishing the tasks within two scripts shown in Figures D.1 to D.8. The scripts have similar steps with the only difference being in choosing which interface of browsing indicators to use at the beginning of the experiment. Script A asks users to choose Tabular view first, while script B asks them to choose Circular view first.
The Circles of Sustainability Software Instructions – Version A

1. You have already received the invitation email. Please first follow the survey link provided in the email.

2. Answer questions 1 to 5 on the first page from the survey about assessing your knowledge of the concepts engaged in this user-study.

3. Next, go to the CoS webpage on http://192.184.92.212

4. You are in the main page now. In Getting started section, please register first. (See Figure 1)

5. After login, you are directed to the Project Listing page. From the Current Projects section, choose Tehran Air Pollution.

6. You are now on the Project Dashboard Page. There are five main sections. (See Figure 2). You can explore each section and spend some time to know about it.
   - **About the Tehran Air Pollution project:** Description of the project and details about City of Tehran and its pollution problem.
   - **Project Progress:** The tasks that are completed or needed to be completed
   - **Common Indicator Sets:** List of existing indicator sets that can be used.
   - **Critical Issues:** List of critical issues that are pre-defined by the experts.
   - **Assessments:** Project assessment that should be done after adding indicators to each issue.

---

Figure 1: Main Page

Welcome to *Circles of Sustainability* assessment tool.

Figure D.1: CoS User Study - Script A - Page 1
7. On the Project Dashboard page, a list of common indicator sets are displayed and two critical issues are already defined (See Figure 3). You can add a new indicator or choose from an existing indicator. It is recommended to choose from existing indicators.

Two tasks are expected from users to accomplish this experiment:

a) Adding relevant indicators to issues.

b) Assessing the project using the added indicators against Circles of Sustainability.

7.1. Click on the first issue. You will be directed to the Issue page that has a description. Read the description and save it. You are now on the Project Dashboard page. For the same issue, select Adding an existing indicator.

7.2. On the Browsing Indicator page, Select Table view (See Figure 3). Table view illustrates the information about indicator in 8 columns. There is also a search option for querying keywords related to the topic of the issue. Once you find the relevant indicator(s) to that issue, click on the label Add Indicator to Issue in the Action column. You will be directed to the Adding page. Save the issue and you will be back to the Project Dashboard page.

7.3. You can repeat step 7.2. and it is possible to add as much as indicators to the issues on the condition of choosing an indicator only once.
8.1. Do the same process for the second issue by repeating steps 7.1 and 7.3 on the condition of selecting Circular view this time. (See Figure 4) This view displays indicators in a circle divided into portions of four domains and 7 subdomains. By clicking on each portion the name of the subdomain is appeared on the left corner of the page. A list of related indicators and their information is also shown on the right side of the page. In this view, reading subdomains is a good way to find relevant indicators. Once you find the relevant indicator(s) to that issue, click on the label Add Indicator to Issue. You will be directed to the Adding page. Save the issue and you will be back to the Project Dashboard page.

8.2. You can repeat step 8.1, and it is possible to add as much as indicators to the issues on the condition of choosing an indicator only once.
9. The second task of the experiment is assessing the project against the Circle of Sustainability. On the Project Dashboard page, click on the Add a New Assessment button. You will be directed to the Assessments page.

10. There are two kinds of indicators you can use as part of your project assessment. Firstly, you can assess each of the Circles of Sustainability subdomains (See Figure 4), by using the panel on the left. Secondly, you can enter values for each of the indicators you have chosen, in the panel on the right. Both the subdomain and indicator values form part of an overall assessment of your project, at a particular point in time.

11. Once you have completed both forms of assessment, click the “Save this Assessment” button. You can add as much as assessment to the project.

12. You have completed the experiment. Please go back to the survey page and answer questions 6 to 25.
The Circles of Sustainability Software Instructions – Version B

1. You have already received the invitation email. Please first follow the survey link provided in the email.

2. Answer questions 1 to 5 on the first page from the survey about assessing your knowledge of the concepts engaged in this user-study.

3. Next, go to the CoS webpage on http://192.184.92.212

4. You are in the main page now. In Getting started section, please register first. (See Figure 1)

5. After login, you are directed to the Project Listing page. From the Current Projects section, choose Tehran Air Pollution.

6. You are now on the Project Dashboard Page. There are five main sections. (See Figure 2): You can explore each section and spend some time to know about it.
   - **About the Tehran Air Pollution project:** Description of the project and details about City of Tehran and its pollution problem.
   - **Project Progress:** The tasks that are completed or needed to be completed
   - **Common Indicator Sets:** List of existing indicator sets that can be used.
   - **Critical Issues:** List of critical issues that are pre-defined by the experts.
   - **Assessments:** Project assessment that should be done after adding indicators to each issue.

---

**Figure D.5: CoS User Study - Script B - Page 1**
7. On the Project Dashboard page, a list of common indicator sets are displayed and two critical issues are already defined (See Figure 3). You can add a new indicator or choose from an existing indicator. It is recommended to choose from existing indicators.

Two tasks are expected from users to accomplish this experiment:

a) Adding relevant indicators to issues.

b) Assessing the project using the added indicators against Circles of Sustainability.

7.1. Click on the first issue. You will be directed to the Issue page that has a description. Read the description and save it. You are now on the Project Dashboard page. For the same issue, select Adding an existing indicator.

7.2. On the Browsing Indicator page, Select Circular view (See Figure 3). This view displays indicators in a circle divided into portions of four domains and 7 subdomains. By clicking on each portion the name of the subdomain is appeared on the left corner of the page. A list of related indicators and their information is also shown on the right side of the page. In this view, reading subdomains is a good way to find relevant indicators. Once you find the relevant indicator(s) to that issue, click on the label Add Indicator to Issue. You will be directed to the Adding page. Save the issue and you will be back to the Project Dashboard page.

7.3. You can repeat step 7.2. and it is possible to add as much as indicators to the issues on the condition of choosing an indicator only once.
8.1. Do the same process for the second issue by repeating steps 7.1 and 7.3 on the condition of selecting Table view this time. (See Figure 4) Table view illustrates the information about indicator in 8 columns. There is also a search option for querying keywords related to the topic of the issue. Once you find the relevant indicator(s) to that issue, click on the label Add Indicator to Issue in the Action column. You will be directed to the Adding page. Save the issue and you will be back to the Project Dashboard page.

8.2. You can repeat step 8.1, and it is possible to add as much as indicators to the issues on the condition of choosing an indicator only once.
9. The second task of the experiment is assessing the project against the Circle of Sustainability. On the Project Dashboard page, click on the Add a New Assessment button. You will be directed to the Assessments page.

10. There are two kinds of indicators you can use as part of your project assessment. 
    Firstly, you can assess each of the Circles of Sustainability subdomains (See Figure 4), by using the panel on the left. Secondly, you can enter values for each of the indicators you have chosen, in the panel on the right. Both the subdomain and indicator values form part of an overall assessment of your project, at a particular point in time.

11. Once you have completed both forms of assessment, click the "Save this Assessment" button. You can add as much as assessment to the project.

12. You have completed the experiment. Please go back to the survey page and answer questions 6 to 25.
Appendix E

CoS User Study - Qualitative Responses

This section provides some of the participants’ responses to the two open questions (Q12 and Q25) and the general statement about the CoS software shown in Figures E.1 to E.4. However, not all participants answered these questions.
Q12. Can you imagine using the software in everyday sustainability assessment and reporting practices? What could be done or added to make it more useful?

The software would be helpful for overall assessment and reporting. If the software could generate assessment “tools” (e.g. ratings scales, survey questions etc.) this would be very useful.

It was unclear what sort of assessment values to enter in the final screen; there is a great deal of variation between the indicators. For example, some are measured as percentages (%), others in volume (e.g. emissions particles), others as “scores.” This makes it difficult to use a single system of reporting or assessment.

Sustainability assessment is not something I do on a regular basis, so it’s difficult for me to comment here. The first task on the experiment was a bit confusing, but I think this is mainly because the software instructions provided could be clearer (there are some grammatical errors and explanatory errors that make it a little confusing in some parts). The instructions might also benefit from some more context and explanation regarding the purpose of task 1. Task 2 was clearer and the saving tool looks really useful and user-friendly.

At the moment the software is uncomfortable to use. I can see the amazing potential of the method, and I think that you are going in the right direction, but many of the steps are awkward and unresolved.

Having a delete button so prominent at the entry point is strange. I did not understand that some of the subdomains would be automatically chosen when I wrote the terms of the issue. The subdomain names appear in a less than obvious way when the cursor is moved onto the subdomain. The system did not seem to stick consistently, and I had to redo some writing.

Based on the level of knowledge, defining different sustainability indicators would be useful, i.e. for experts the indicators can be designed so detailed and professional and for non-professional people the indicators of the same assessment should be designed simply, i.e. the low level indicators and high level indicators uniform to a unique indicator based on their relational weight.

The COS software would definitely assist and support the assessment and reporting practices and process as it gives an immediate, tangible and readable results through the assessments.

I would suggest a more visually appealing graphic design of the software.

As somebody who is not involved in sustainability assessment I found the software a little difficult to navigate. The circle view is very helpful. The table view was quite hard to figure out.

Software for CoS would take more effort when compared with another software.

Figure E.1: Participants’ Responses to Q12
Q25. The software ultimately aims to help an organisation to select indicators for measuring a series of critical issues that a stakeholder group has identified. In addressing this goal, which view (Table or Circular view) do you prefer most? Can you briefly explain your reasons.

- Circular, better representation

- Table view, if it would be more flexible.

- Table view, because it is more user-friendly and organized.

- I prefer the circular view. I think it is visually easier to navigate.

- I would prefer to use table view as it represents the environment in a more holistic and realistic manner.

- Both views are useful, but the method of selecting the indicators is still unwieldy. A large part of the problem is that there are not enough relevant indicators in the system. In the table view browsing is slow, and a keyword search did not throw up the kind of indicators that I thought would be useful. I can see how it would be magnificent and save a lot of time if the system was fully populated with dozens of relevant indicators in each subdomain.

- I think both views can be interchangeable by software. Tables vs Circle view depends upon the number of indicators. For fewer variety of indicators, circular model would be ok but for complicated models with numerous indicators the tabular would be more efficient and usable.

- Table view easier to browse among indicators and compare them. Circular view is better for a graphic display which then could be use for reporting more easily. I would leave both options available for the preference of the user.

Figure E.2: Participants’ Responses to Q25 - Part 1
Figure E.3: Participants’ Responses to Q25 - Part 2

Q25. The software ultimately aims to help an organisation to select indicators for measuring a series of critical issues that a stakeholder group has identified. In addressing this goal, which view (Table or Circular view) do you prefer most? Can you briefly explain your reasons.

The circular view is a visual reflection of the table and provides a comparative visual experience of the assessment process.

I prefer circular view because it makes me perceive results in a quicker and more pleasant way. It would be useful to have the option to select additional areas of information per each section of the circle by simply clicking with the mouse on a single segment of the circle.

I prefer the circular view because I could understand it. There was too much text involved with the table view, which made it difficult to obtain an overall picture of the different indicators.

I would prefer Table because there are many text related to each indicator and would be useful to view them all together at one. Circle also would be useful to give a general idea and short details.

I prefer table view as it is easier to see a number of different indicators simultaneously. This is important for comparison of indicators - as some are very similar. Also indicators can be viewed according to the individual frameworks e.g. OECD.

Circular view is far easier to use because it only requires the user's intuition to figure out. Each piece of the circle can easily be identified and manipulated. In contrast, the table view is less visually oriented and more complex to use. Its exactness makes it ideal for reporting specific values, while simultaneously challenging to understand during the first use. The circular view may be more appealing for users with a language barrier.

Table view. Personal bias, I like to have all the information in one searchable table

Table view, if it would be more flexible.
. Please let us know any other feedback you have about the software.

The circles inside the process would be better having the subdomains named directly against them or coming up with spatial contiguity with the cursor.

The software needs more instructions about what it is doing.

The 'Add a new issue' button needs to be separated off from the 'Add a new indicator'

The software can be designed in a way that if in a project we use the tabular model, the circular model generated automatically. Exporting the tabular model data to spreadsheets would be useful as well.

Delete buttons in not catch the attention over the rest, which can be risky.

I accidently clicked the delete button in project listings due to color and because there was no other button such as "go" or "continue".

Possibly the process may be facilitated and make it more user friendly through the dashboard selection. Personally found it difficult to erase an assessment once logged.

I noticed that when I want to add indicators to one of the two 'critical issues', I am given a choice only in the range of GRI Indicator but, while the indicators of UN Habitat Agenda and OECD are not accessible. So, in this experiment, I did not have a proper choice for the issue N. 1.

I think this software could more helpful for domain expert more than for public users.

In table view it would be helpful if indicators could be "refined" and sorted based on criteria such as "tags", keywords, themes etc.

Figure E.4: Participants’ Responses to General Statement
Appendix F

Glossary

AHP Analytic Hierarchy Process
AI Artificial Intelligence
ARC Australian Research Council
CMC Characterising Computer-mediated Communication
CoS Circles of Sustainability
CPs Content ODPs
DBMS Database Management System
DL Description Logics
DOLCE Descriptive Ontology for Linguistic and Cognitive Engineering
ESA Emergency Situation Awareness
GOSIS Generic Ontology for Sustainability Indicator Sets
GRI Global Reporting Initiative
HCI Human Computer Interaction
IE Information Extraction
IR Information Retrieval
ISDs Indicators and Sustainable Development
KBSs Knowledge-Based Systems
MTC Multi-level Tree of Characteristics
MVC Model-View-Controller
NER Named-Entity Recognition
NGOs Non-government Organisations
NLP Natural Language Processing
ODP Open Directory Project
ODPs Ontology design patterns
OECD Organization for Economic Co-operation and Development
OSIS Ontology for Sustainability Indicator Sets
OWL Web Ontology Language
OzCT OzCrisisTracker
PNS Professional Networking Services
POS Part of Speech
PSMs Problem Solving Methods
RDF Resource Description Framework
RDF-S RDF Schema
ROMEO Requirements Oriented Methodology for Evaluating Ontologies
RT ReTweet

SDPs Software Design Patterns

SIOC Semantically Interlinked Online Communities

SNS Social Network Sites

SOSIS Specific Ontology for Sustainability Indicator Sets

SUMO Suggested Upper Merged Ontology

SUO Standard Upper Ontology

SVM Support Vector Machine

UN United Nations

URI Uniform Resource Identifier

VP Value Partition

W3C World Wide Web Consortium

XBRL eXtensible Business Reporting Language

XML Extensible Markup Language
Bibliography


B. R. Gaines, M. L. Shaw, and J. B. Woodward. Modeling as framework for knowledge


